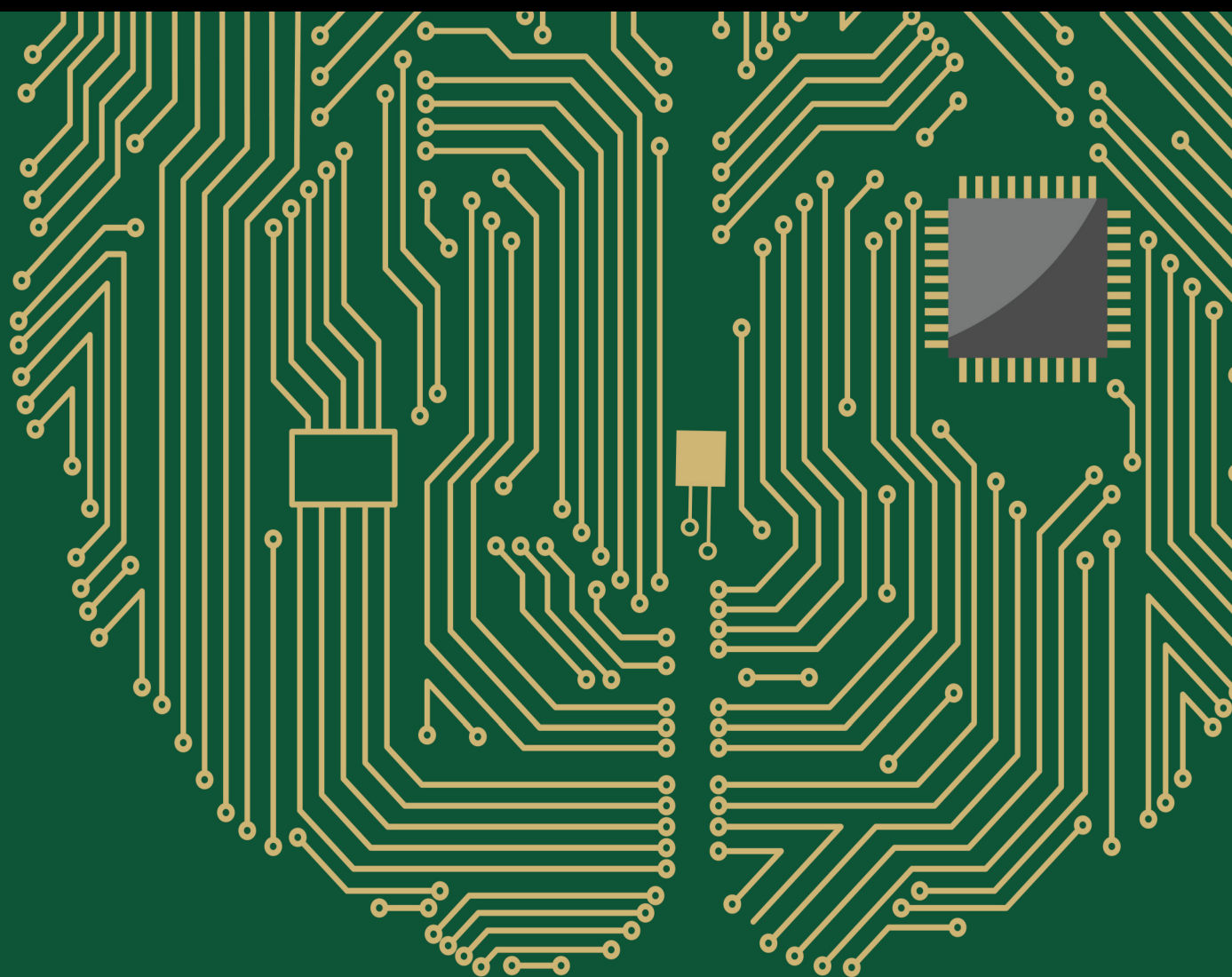


# Multi-Objective Intelligent Decision-Making Methods Driven by Big Data

Lead Guest Editor: Daqing Gong

Guest Editors: Aboul Ella Hassanien and Jinan Fiaidhi







---

# **Multi-Objective Intelligent Decision-Making Methods Driven by Big Data**

Computational Intelligence and Neuroscience

---

## **Multi-Objective Intelligent Decision-Making Methods Driven by Big Data**

Lead Guest Editor: Daqing Gong

Guest Editors: Aboul Ella Hassanien and Jinan  
Fiaidhi



Copyright © 2023 Hindawi Limited. All rights reserved.

This is a special issue published in "Computational Intelligence and Neuroscience." All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

# Chief Editor

Andrzej Cichocki, Poland

## Associate Editors

Arnaud Delorme, France  
Cheng-Jian Lin , Taiwan  
Saeid Sanei, United Kingdom

## Academic Editors

Mohamed Abd Elaziz , Egypt  
Tariq Ahanger , Saudi Arabia  
Muhammad Ahmad, Pakistan  
Ricardo Aler , Spain  
Nouman Ali, Pakistan  
Pietro Aricò , Italy  
Lerina Aversano , Italy  
Ümit Ağbulut , Turkey  
Najib Ben Aoun , Saudi Arabia  
Surbhi Bhatia , Saudi Arabia  
Daniele Bibbo , Italy  
Vince D. Calhoun , USA  
Francesco Camastra, Italy  
Zhicheng Cao, China  
Hubert Cecotti , USA  
Jyotir Moy Chatterjee , Nepal  
Rupesh Chikara, USA  
Marta Cimitile, Italy  
Silvia Conforto , Italy  
Paolo Crippa , Italy  
Christian W. Dawson, United Kingdom  
Carmen De Maio , Italy  
Thomas DeMarse , USA  
Maria Jose Del Jesus, Spain  
Arnaud Delorme , France  
Anastasios D. Doulamis, Greece  
António Dourado , Portugal  
Sheng Du , China  
Said El Kafhali , Morocco  
Mohammad Reza Feizi Derakhshi , Iran  
Quanxi Feng, China  
Zhong-kai Feng, China  
Steven L. Fernandes, USA  
Agostino Forestiero , Italy  
Piotr Franaszczuk , USA  
Thippa Reddy Gadekallu , India  
Paolo Gastaldo , Italy  
Samanwoy Ghosh-Dastidar, USA

Manuel Graña , Spain  
Alberto Guillén , Spain  
Gaurav Gupta, India  
Rodolfo E. Haber , Spain  
Usman Habib , Pakistan  
Anandakumar Haldorai , India  
José Alfredo Hernández-Pérez , Mexico  
Luis Javier Herrera , Spain  
Alexander Hošovský , Slovakia  
Etienne Hugues, USA  
Nadeem Iqbal , Pakistan  
Sajad Jafari, Iran  
Abdul Rehman Javed , Pakistan  
Jing Jin , China  
Li Jin, United Kingdom  
Kanak Kalita, India  
Ryotaro Kamimura , Japan  
Pasi A. Karjalainen , Finland  
Anitha Karthikeyan, Saint Vincent and the Grenadines  
Elpida Keravnou , Cyprus  
Asif Irshad Khan , Saudi Arabia  
Muhammad Adnan Khan , Republic of Korea  
Abbas Khosravi, Australia  
Tai-hoon Kim, Republic of Korea  
Li-Wei Ko , Taiwan  
Raşit Köker , Turkey  
Deepika Koundal , India  
Sunil Kumar , India  
Fabio La Foresta, Italy  
Kuruva Lakshmana , India  
Maciej Lawrynczuk , Poland  
Jianli Liu , China  
Giosuè Lo Bosco , Italy  
Andrea Loddo , Italy  
Kezhi Mao, Singapore  
Paolo Massobrio , Italy  
Gerard McKee, Nigeria  
Mohit Mittal , France  
Paulo Moura Oliveira , Portugal  
Debajyoti Mukhopadhyay , India  
Xin Ning , China  
Nasimul Noman , Australia  
Fivos Panetsos , Spain

Evgeniya Pankratova , Russia  
Rocío Pérez de Prado , Spain  
Francesco Pistolesi , Italy  
Alessandro Sebastian Podda , Italy  
David M Powers, Australia  
Radu-Emil Precup, Romania  
Lorenzo Putzu, Italy  
S P Raja, India  
Dr.Anand Singh Rajawat , India  
Simone Ranaldi , Italy  
Upaka Rathnayake, Sri Lanka  
Navid Razmjoo, Iran  
Carlo Ricciardi, Italy  
Jatinderkumar R. Saini , India  
Sandhya Samarasinghe , New Zealand  
Friedhelm Schwenker, Germany  
Mijanur Rahaman Seikh, India  
Tapan Senapati , China  
Mohammed Shuaib , Malaysia  
Kamran Siddique , USA  
Gaurav Singal, India  
Akansha Singh , India  
Chiranjibi Sitaula , Australia  
Neelakandan Subramani, India  
Le Sun, China  
Rawia Tahrir , Iraq  
Binhua Tang , China  
Carlos M. Travieso-González , Spain  
Vinh Truong Hoang , Vietnam  
Fath U Min Ullah , Republic of Korea  
Pablo Varona , Spain  
Roberto A. Vazquez , Mexico  
Mario Versaci, Italy  
Gennaro Vessio , Italy  
Ivan Volosyak , Germany  
Leyi Wei , China  
Jianghui Wen, China  
Lingwei Xu , China  
Cornelio Yáñez-Márquez, Mexico  
Zaher Mundher Yaseen, Iraq  
Yugen Yi , China  
Qiangqiang Yuan , China  
Miaolei Zhou , China  
Michal Zochowski, USA  
Rodolfo Zunino, Italy



# Contents

---

**Retracted: A Data-Driven Control Strategy for Urban Express Ramp**

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9761013, Volume 2023 (2023)

**Retracted: Modeling on Social Health Literacy Level Prediction**

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9780743, Volume 2023 (2023)

**Retracted: Security and Privacy Risk Assessment of Energy Big Data in Cloud Environment**

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9896475, Volume 2023 (2023)

**Retracted: Intelligent Control of Agricultural Irrigation through Water Demand Prediction Based on Artificial Neural Network**

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9878132, Volume 2023 (2023)

**Retracted: A DSGE Decision Model for Investigating the LPR Transmission Effect**

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9878102, Volume 2023 (2023)

**Retracted: Optimal Defect Detection and Sensing System of Railway Tunnel Radar considering Multisensor System Combined with Active Interference Suppression Algorithm**

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9836784, Volume 2023 (2023)

**Retracted: Development of a Relative Similarity Degree Based Engineering Construction Multi-Attribute Decision Model and Its Application**

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9834689, Volume 2023 (2023)

**Retracted: Cloud Computing Load Balancing Mechanism Taking into Account Load Balancing Ant Colony Optimization Algorithm**

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9831926, Volume 2023 (2023)

**Retracted: A Decision Method for Benefit Distribution Mechanism of Shared End Distribution on Shapley Value**

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9819586, Volume 2023 (2023)

**Retracted: Supply Chain Decision Analysis of Community E-Commerce Platform under Different Power Structures: Considering the Influence of Value Cocreation**

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9815690, Volume 2023 (2023)

**Retracted: Local Privacy Protection for Sensitive Areas in Multiface Images**

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9898736, Volume 2023 (2023)

**Retracted: Generalized Zero-Adjusted Models to Predict Medical Expenditures**

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9893514, Volume 2023 (2023)

**Retracted: An Optimized Decision Method for Smart Teaching Effect Based on Cloud Computing and Deep Learning**

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9862737, Volume 2023 (2023)

**Retracted: Innovative Research on Urban Community Governance Decision-Making Relying on Distributed High-Performance Computing Blockchain Key Algorithms**

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9847573, Volume 2023 (2023)

**Retracted: An Individual Driving Behavior Portrait Approach for Professional Driver of HDVs with Naturalistic Driving Data**

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9830541, Volume 2023 (2023)

**Retracted: A Double-Layered Belief Rule Base Model for Driving Anger Detection Using Human, Vehicle, and Environment Characteristics: A Naturalistic Experimental Study**

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9793263, Volume 2023 (2023)

**Retracted: Sliding Mode Control of Flexible Articulated Manipulator Based on Robust Observer**

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9826408, Volume 2023 (2023)

**Retracted: Design of Sponsored Search Auction Mechanism for Federated Learning Advertising Platform**

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9892141, Volume 2023 (2023)

**Retracted: Tripartite Evolutionary Game Analysis of Financial Support for Science-Tech Enterprises' Innovation with Government Participation**

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9852672, Volume 2023 (2023)

**Retracted: A Study on the Impact of Regional Total Factor Production in Digital Economy Based on Fuzzy Hierarchical VISC Algorithm**

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9836743, Volume 2023 (2023)

# Contents

---

**Retracted: Managing the Gate Assignment Problem in the Hub Airport with Satellite Halls: A Transfer Demand-Oriented Approach**

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9820329, Volume 2023 (2023)

**Retracted: An Algorithm for Time Prediction Signal Interference Detection Based on the LSTM-SVM Model**

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9805804, Volume 2023 (2023)

**Retracted: A Novel Hybrid Algorithm for Multiobjective Location-Allocation Problem in Emergency Logistics**

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9764054, Volume 2023 (2023)

**Retracted: A Prediction Model Analysis of Behavior Recognition Based on Genetic Algorithm and Neural Network**

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9840187, Volume 2023 (2023)

**Retracted: Multirobot Collaborative Pursuit Target Robot by Improved MADDPG**

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9839345, Volume 2023 (2023)

**Retracted: Modeling and Optimization of Water Distribution in Mineral Processing considering Water Cost and Recycled Water**

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9838763, Volume 2023 (2023)

**Retracted: Brand Marketing Leveraging the Advantage of Emoji Pack Relying on Association Rule Algorithm in Data Mining Technology**

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9835685, Volume 2023 (2023)

**Retracted: Dynamic Mechanism and Decision Analysis of Urban Exhibition Tourism Development Based on High-Performance Computing**

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9831376, Volume 2023 (2023)

**Retracted: Picture Fuzzy Einstein Hybrid-Weighted Aggregation Operator and Its Application to Multicriteria Group Decision Making**

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9826814, Volume 2023 (2023)

**Retracted: An Annealing Model Analysis and Research on the In of Community Public Sports on Economic Development**

Computational Intelligence and Neuroscience  
Retraction (1 page), Article ID 9826208, Volume 2023 (2023)

**Retracted: Novel Decision-Making Techniques in Tripolar Fuzzy Environment with Application: A Case Study of ERP Systems**

Computational Intelligence and Neuroscience  
Retraction (1 page), Article ID 9819486, Volume 2023 (2023)

**Retracted: Random Forest Algorithm Based on Speech for Early Identification of Parkinson's Disease**

Computational Intelligence and Neuroscience  
Retraction (1 page), Article ID 9796715, Volume 2023 (2023)

**Retracted: A Cost Allocation Decision Model for Air Pollution Control**

Computational Intelligence and Neuroscience  
Retraction (1 page), Article ID 9794657, Volume 2023 (2023)

**Retracted: Cluster-Based Mutual Fund Classification and Price Prediction Using Machine Learning for Robo-Advisors**

Computational Intelligence and Neuroscience  
Retraction (1 page), Article ID 9789518, Volume 2023 (2023)

**Retracted: Application of Mathematical Models in Economic Variable Input and Output Models under the Scientific Visualization**

Computational Intelligence and Neuroscience  
Retraction (1 page), Article ID 9785691, Volume 2023 (2023)

**Retracted: RFID Data Analysis and Evaluation Based on Big Data and Data Clustering**

Computational Intelligence and Neuroscience  
Retraction (1 page), Article ID 9780823, Volume 2023 (2023)

**Retracted: Visualization and Analysis of Mapping Knowledge Domain of Heterogeneous Traffic Flow**

Computational Intelligence and Neuroscience  
Retraction (1 page), Article ID 9769504, Volume 2023 (2023)

**Retracted: A Graph Partition-Based Large-Scale Distribution Network Reconfiguration Method**

Computational Intelligence and Neuroscience  
Retraction (1 page), Article ID 9769308, Volume 2023 (2023)

**Retracted: Research and Design of Docker Technology Based Authority Management System**

Computational Intelligence and Neuroscience  
Retraction (1 page), Article ID 9767210, Volume 2023 (2023)

# Contents

---

**Retracted: Differentially Private Singular Value Decomposition for Training Support Vector Machines**

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9760375, Volume 2023 (2023)

**Retracted: A Mode of Intelligent Equipment Monitoring Optimization Based on Dynamic Programming Algorithm**

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9758926, Volume 2023 (2023)

**Retracted: Digital Protection and Development of Intangible Cultural Heritage Relying on High-Performance Computing**

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9756258, Volume 2023 (2023)

**Retracted: A Game Model for Incentive Mechanism of Distributed Nodes in Supply Chains**

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9893750, Volume 2023 (2023)

**Retracted: High-Performance Computational Recognition of Communication Signals Based on Bispectral Quadratic Feature Model**

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9873784, Volume 2023 (2023)

**Retracted: Research on Probability Distribution of Short-Term Photovoltaic Output Forecast Error Based on Numerical Characteristic Clustering**

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9863081, Volume 2023 (2023)

**Retracted: Computational Prediction of Intrinsically Disordered Proteins Based on Protein Sequences and Convolutional Neural Networks**

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9861582, Volume 2023 (2023)

**Retracted: Order Management and Completion Date Prediction of Manufacturing Job-Shop Based on Deep Learning**

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9860295, Volume 2023 (2023)

**Retracted: Spare Parts Stocking Decision Strategy and Service Logistics Cost Optimization of Two-Echelon Service Logistics System considering Multifailure Mode**

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9854834, Volume 2023 (2023)

**Retracted: A Take-Over Performance Evaluation Model for Automated Vehicles from Automated to Manual Driving**

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9848737, Volume 2023 (2023)



**Retracted: Feature Sequencing Method of Industrial Control Data Set Based on Multidimensional Evaluation Parameters**

Computational Intelligence and Neuroscience  
Retraction (1 page), Article ID 9847359, Volume 2023 (2023)

**Retracted: A Prediction Model of Health Development Based on Linear Sequential Extreme Learning Machine Algorithm Matrix**

Computational Intelligence and Neuroscience  
Retraction (1 page), Article ID 9847150, Volume 2023 (2023)


**Retracted: Discontinuous Track Recognition System Based on PolyLaneNet for Darwin-op2 Robot**

Computational Intelligence and Neuroscience  
Retraction (1 page), Article ID 9842798, Volume 2023 (2023)


**Retracted: Dynamic Performance Analysis of STEP System in Internet of Vehicles Based on Queuing Theory**

Computational Intelligence and Neuroscience  
Retraction (1 page), Article ID 9842681, Volume 2023 (2023)





**[Retracted] Development of a Relative Similarity Degree Based Engineering Construction Multi-Attribute Decision Model and Its Application**

Ling Sui, Xiaojie Liang , and Qilong Huang  
Research Article (5 pages), Article ID 4274795, Volume 2022 (2022)


**[Retracted] Brand Marketing Leveraging the Advantage of Emoji Pack Relying on Association Rule Algorithm in Data Mining Technology**

Huan Tian   
Research Article (9 pages), Article ID 3511211, Volume 2022 (2022)


**A Novel U-Net Based Deep Learning Method for 3D Cardiovascular MRI Segmentation**

Yinan Lu , Yan Zhao , Xing Chen , and Xiaoxin Guo   
Research Article (11 pages), Article ID 4103524, Volume 2022 (2022)

**[Retracted] Picture Fuzzy Einstein Hybrid-Weighted Aggregation Operator and Its Application to Multicriteria Group Decision Making**

Guo Cao   
Research Article (20 pages), Article ID 6925670, Volume 2022 (2022)

**[Retracted] Random Forest Algorithm Based on Speech for Early Identification of Parkinson's Disease**




Ping Fan   
Research Article (6 pages), Article ID 3287068, Volume 2022 (2022)

**Automatic Rule Generation for Decision-Making in Context-Aware Systems Using Machine Learning**

Roua Jabla , Maha Khemaja , Félix Buendia , and Sami Faiz   
Research Article (13 pages), Article ID 5202537, Volume 2022 (2022)

# Contents

## **An Efficient Data Classification Decision Based on Multimodel Deep Learning**

Wenjin Hu , Feng Liu , and Jiebo Peng 

Research Article (10 pages), Article ID 7636705, Volume 2022 (2022)

## **[Retracted] Tripartite Evolutionary Game Analysis of Financial Support for Science-Tech Enterprises' Innovation with Government Participation**

Wei Liu , Yang Liu , Jun Gu , Jing Zhao , and Rong Zhang 


Research Article (21 pages), Article ID 5207003, Volume 2022 (2022)

## **[Retracted] Feature Sequencing Method of Industrial Control Data Set Based on Multidimensional Evaluation Parameters**

Xue-Jun Liu , Xiang-Min Kong , Xiao-Ni Zhang , Hai-Ying Luan, Yong Yan , Yun Sha , Kai-Li Li , Xue-Ying Cao , and Jian-Ping Chen 


Research Article (10 pages), Article ID 9248267, Volume 2022 (2022)

## **[Retracted] High-Performance Computational Recognition of Communication Signals Based on Bispectral Quadratic Feature Model**

Yarong Chen , Rui Zhu, Jianxin Guo, and Feng Wang

Research Article (12 pages), Article ID 2773492, Volume 2022 (2022)

## **[Retracted] Managing the Gate Assignment Problem in the Hub Airport with Satellite Halls: A Transfer Demand-Oriented Approach**

Runfa Wu , Wenliang Zhou , Linhuan Zhong , and Yihan Liu 






Research Article (15 pages), Article ID 5087941, Volume 2022 (2022)

## **[Retracted] Optimal Defect Detection and Sensing System of Railway Tunnel Radar considering Multisensor System Combined with Active Interference Suppression Algorithm**

Yang Lei , Yong Zou , Bo Jiang, and Tian Tian






Research Article (13 pages), Article ID 2459996, Volume 2022 (2022)

## **Fault Diagnosis in Regenerative Braking System of Hybrid Electric Vehicles by Using Semigroup of Finite-State Deterministic Fully Intuitionistic Fuzzy Automata**

Sajida Kousar , Farah Aslam , Nasreen Kausar , Dragan Pamucar , and Gezahagne Mulat Addis 






Research Article (14 pages), Article ID 3684727, Volume 2022 (2022)

## **Cross-Project Defect Prediction Based on Two-Phase Feature Importance Amplification**

Ying Xing , Wanting Lin , Xueyan Lin , Bin Yang , and Zhou Tan 

Research Article (14 pages), Article ID 2320447, Volume 2022 (2022)

## **Combined Economic Emission Dispatch of Microgrid with the Incorporation of Renewable Energy Sources Using Improved Mayfly Optimization Algorithm**

Karthik Nagarajan , Arul Rajagopalan , S. Angalaeswari , L. Natrayan , and Wubishet Degife Mammo 

Research Article (22 pages), Article ID 6461690, Volume 2022 (2022)

**[Retracted] A Take-Over Performance Evaluation Model for Automated Vehicles from Automated to Manual Driving**

Lixin Yan , Jiayu Chen , Chengyue Wen , Ping Wan , Liqun Peng , and Xujin Yu 


Research Article (16 pages), Article ID 3160449, Volume 2022 (2022)

**[Retracted] An Annealing Model Analysis and Research on the In of Community Public Sports on Economic Development**

Zhen Guo 

Research Article (9 pages), Article ID 3129638, Volume 2022 (2022)

**[Retracted] Research and Design of Docker Technology Based Authority Management System**

Li You and Hui Sun 


Research Article (8 pages), Article ID 5325694, Volume 2022 (2022)

**[Retracted] Digital Protection and Development of Intangible Cultural Heritage Relying on High-Performance Computing**

Hong Ding 


Research Article (10 pages), Article ID 4955380, Volume 2022 (2022)

**[Retracted] Cloud Computing Load Balancing Mechanism Taking into Account Load Balancing Ant Colony Optimization Algorithm**

Jing He 


Research Article (10 pages), Article ID 3120883, Volume 2022 (2022)

**Multidimensional Information Network Big Data Mining Algorithm Relying on Finite Element Analysis**

Haifeng Li 




Research Article (11 pages), Article ID 7156715, Volume 2022 (2022)

**Empirical Study of Large-Scale HLA Simulation of Parallel Region-Matching Knowledge Recognition Algorithm Based on Region Matching**

Guohua Zhu  and Haizhou Wang



Research Article (9 pages), Article ID 1514396, Volume 2022 (2022)

**[Retracted] Dynamic Performance Analysis of STEP System in Internet of Vehicles Based on Queuing Theory**

Huiyong Li , Xiaofeng Wu , and Yanhong Wang 

Research Article (13 pages), Article ID 8322029, Volume 2022 (2022)


**[Retracted] Design of Sponsored Search Auction Mechanism for Federated Learning Advertising Platform**

Hong Jiang , Tianxu Cui , and Kaiwen Yang

Research Article (15 pages), Article ID 5787491, Volume 2022 (2022)

# Contents

**[Retracted] Dynamic Mechanism and Decision Analysis of Urban Exhibition Tourism Development Based on High-Performance Computing**

Hongzhen Wan 



Research Article (14 pages), Article ID 3806752, Volume 2022 (2022)

**Research on Subsidy Strategy of Shared Accommodation Platform under the Background of Big Data Based on Evolutionary Game**

Meng Xiao  and Chang Zhai



Research Article (22 pages), Article ID 9066215, Volume 2022 (2022)

**Lightweight YOLOv4 with Multiple Receptive Fields for Detection of Pulmonary Tuberculosis**

Zhitao Guo , Jiahao Wang, Jinghua Wang, and Jinli Yuan 


Research Article (13 pages), Article ID 9465646, Volume 2022 (2022)

**Simulation Study on Complex Systems of Forest Biomass Power Generation Supply Chain in China**

Ning Ma , Zhao Huang , and Yue Qi 


Research Article (14 pages), Article ID 7202352, Volume 2022 (2022)

**Exploratory Research on Key Technology of Human-Computer Interactive 2.5-Minute Fast Digital Early Warning for Mild Cognitive Impairment**

Nan Li, Xiaotong Yang, Wencai Du, Atsushi Ogihara, Siyu Zhou, Xiaowen Ma, Yujia Wang, Shuwu Li, and Kai Li 

Research Article (15 pages), Article ID 2495330, Volume 2022 (2022)

**[Retracted] RFID Data Analysis and Evaluation Based on Big Data and Data Clustering**

Lihua Lv 


Research Article (10 pages), Article ID 3432688, Volume 2022 (2022)

**[Retracted] An Optimized Decision Method for Smart Teaching Effect Based on Cloud Computing and Deep Learning**

Miaomiao Jiang  and Yuwei Sun 

Research Article (10 pages), Article ID 6907172, Volume 2022 (2022)

**[Retracted] A Prediction Model Analysis of Behavior Recognition Based on Genetic Algorithm and Neural Network**

Qifu Wang and Shuzhi Liu 


Research Article (11 pages), Article ID 3552908, Volume 2022 (2022)

**[Retracted] Differentially Private Singular Value Decomposition for Training Support Vector Machines**

Zhenlong Sun , Jing Yang , and Xiaoye Li 


Research Article (11 pages), Article ID 2935975, Volume 2022 (2022)

**Application of a Hybrid Model of Big Data and BP Network on Fault Diagnosis Strategy for Microgrid**



Chundi Jiang and Zhiliang Xia 

Research Article (12 pages), Article ID 1554422, Volume 2022 (2022)

**Production Task Allocation Decision Based on Cloud Robot Cell-Line**

Yi Du , Yongqiang Wang, Jialin Wang, and Tianpeng Zhang  
Research Article (7 pages), Article ID 5892943, Volume 2022 (2022)

**A Novel Approach for Characterizing Solutions of Rough Optimization Problems Based on Boundary Region**

Hamiden Abd El- Wahed Khalifa , Dragan Pamucar , Amina Hadj Kacem, and W. A. Afifi  
Research Article (12 pages), Article ID 8662289, Volume 2022 (2022)

**Decision Scheduling for Cloud Computing Tasks Relying on Solving Large Linear Systems of Equations**

Jing He   
Research Article (12 pages), Article ID 3411959, Volume 2022 (2022)





**Supervision Strategy Analysis on Price Discrimination of E-Commerce Company in the Context of Big Data Based on Four-Party Evolutionary Game**

Meng Xiao   
Research Article (25 pages), Article ID 2900286, Volume 2022 (2022)

**[Retracted] A Mode of Intelligent Equipment Monitoring Optimization Based on Dynamic Programming Algorithm**

Zhilei Zheng  and Chuan Jiang   
Research Article (11 pages), Article ID 1569428, Volume 2022 (2022)






**A Random Forest Method for Identifying the Effectiveness of Innovation Factor Allocation**

Mo Xu , Yawei Qi , Changqi Tao , and Shangfeng Zhang   
Research Article (16 pages), Article ID 1135582, Volume 2022 (2022)


**[Retracted] Order Management and Completion Date Prediction of Manufacturing Job-Shop Based on Deep Learning**

Mei Wang   
Research Article (10 pages), Article ID 3458116, Volume 2022 (2022)

**Research on the Behavior Evolution of Interest Groups in Book Reverse Logistics**

Shuihai Dou , Xianyang Su , Yanping Du , Yanming Tang , and Lizhi Peng   
Research Article (15 pages), Article ID 3324312, Volume 2022 (2022)

**[Retracted] Innovative Research on Urban Community Governance Decision-Making Relying on Distributed High-Performance Computing Blockchain Key Algorithms**

Jun Wen  and Peihong Xie  
Research Article (13 pages), Article ID 4078014, Volume 2022 (2022)





**A Boosting-Based Deep Distance Metric Learning Method**

Zilong Li   
Research Article (9 pages), Article ID 2665843, Volume 2022 (2022)




# Contents

**[Retracted] Local Privacy Protection for Sensitive Areas in Multiface Images**

Chao Liu , Jing Yang , Xuan Zhang, Yining Zhang , Weinan Zhao , Fengjuan Miao, and Yukun Shao



Research Article (15 pages), Article ID 5919522, Volume 2022 (2022)

**[Retracted] A Graph Partition-Based Large-Scale Distribution Network Reconfiguration Method**

Yuanxia Sha 






Research Article (13 pages), Article ID 3169065, Volume 2022 (2022)

**[Retracted] An Algorithm for Time Prediction Signal Interference Detection Based on the LSTM-SVM Model**

Ningbo Xiao  and Zuxun Song 



Research Article (8 pages), Article ID 1626458, Volume 2022 (2022)

**Carbon Transfer Decision Model Based on LMDI Method**

Yawei Qi , Guangping Rao , Lei Zha , Lu Chen , and Yuping Niu 


Research Article (9 pages), Article ID 3970880, Volume 2022 (2022)

**[Retracted] Spare Parts Stocking Decision Strategy and Service Logistics Cost Optimization of Two-Echelon Service Logistics System considering Multifailure Mode**

Ruiqi Wang  and Guangyu Chen 

Research Article (12 pages), Article ID 7607985, Volume 2022 (2022)

**[Retracted] A Prediction Model of Health Development Based on Linear Sequential Extreme Learning Machine Algorithm Matrix**

Suli Cheng and Shuzhi Liu 

Research Article (12 pages), Article ID 7632841, Volume 2022 (2022)

**Carbon Emission Calculation and Influencing Factor Analysis Based on Industrial Big Data in the “Double Carbon” Era**

Lu Zhang , Yan Yan , Wei Xu , Jun Sun , and Yuanyuan Zhang 


Research Article (12 pages), Article ID 2815940, Volume 2022 (2022)

**A Risk Decision-Making Model Based on Kalman Filter for Intellectual Property Pledge Financing**

Xianan Yin , Hua Ming , and Xinzhong Bao 

Research Article (10 pages), Article ID 8025455, Volume 2022 (2022)

**[Retracted] A Study on the Impact of Regional Total Factor Production in Digital Economy Based on Fuzzy Hierarchical VISC Algorithm**

Haixiang Li and Weixian Xue 

Research Article (11 pages), Article ID 6903836, Volume 2022 (2022)

**Seismic Random Noise Denoising Using Mini-Batch Multivariate Variational Mode Decomposition**

Guoning Wu , Guochang Liu , Junxian Wang, and Pingping Fan


Research Article (14 pages), Article ID 2132732, Volume 2022 (2022)

**[Retracted] Multirobot Collaborative Pursuit Target Robot by Improved MADDPG**

Xiao Zhou , Song Zhou , Xingang Mou , and Yi He 








Research Article (10 pages), Article ID 4757394, Volume 2022 (2022)

**[Retracted] Application of Mathematical Models in Economic Variable Input and Output Models under the Scientific Visualization**

Kai Zheng and Yisheng Liu 

Research Article (10 pages), Article ID 6269358, Volume 2022 (2022)

**Multiobjective Fusion Decision Model for Key Technologies of Aircraft Intelligent Lightweight**

Anliang Zhou , Lei Dong , Yuedong Lang , Jiangping Cui , Wei Ren , Jiankang Liu , and Zhiqiang Wang 


Research Article (8 pages), Article ID 1447718, Volume 2022 (2022)

**[Retracted] A Decision Method for Benefit Distribution Mechanism of Shared End Distribution on Shapley Value**

Hao Zhang , Yuan Hou , and Feng-Feng He 




Research Article (13 pages), Article ID 3816440, Volume 2022 (2022)

**Image Encryption Based on Hopfield Neural Network and Bidirectional Flipping**

Haitao Zhang  and Shuangqi Yang

Research Article (7 pages), Article ID 7941448, Volume 2022 (2022)

**[Retracted] Visualization and Analysis of Mapping Knowledge Domain of Heterogeneous Traffic Flow**

Yi He , Qi Feng , Lixin Yan , and Xiao-Yun Lu

Research Article (15 pages), Article ID 7754961, Volume 2022 (2022)

**An Improved Simulated Annealing-Based Decision Model for the Hybrid Flow Shop Scheduling of Aviation Ordnance Handling**

Xianglei Meng , Nengjian Wang , Jue Liu , and Qinhui Liu 



Research Article (12 pages), Article ID 1843675, Volume 2022 (2022)

**[Retracted] Research on Probability Distribution of Short-Term Photovoltaic Output Forecast Error Based on Numerical Characteristic Clustering**

Peng Yan , Chenmeng Xiang , Tiecheng Li , Xuekai Hu , Wen Zhou , Lei Wang , and Liang Meng 

Research Article (11 pages), Article ID 5355286, Volume 2022 (2022)


**[Retracted] Discontinuous Track Recognition System Based on PolyLaneNet for Darwin-op2 Robot**

Xi-Bao Wu, Si-Chuan Lv, Xiao-Hao Wang, Shi-Zhuo Zhang, Qiong Liu, Yi-Qun Wang , and Wen-bai Chen 

Research Article (10 pages), Article ID 5431886, Volume 2022 (2022)



## Contents

**[Retracted] Modeling and Optimization of Water Distribution in Mineral Processing considering Water Cost and Recycled Water**

He Wang , Qiuyu Cui, and Hualong Du



Research Article (12 pages), Article ID 2314788, Volume 2022 (2022)

**[Retracted] Novel Decision-Making Techniques in Tripolar Fuzzy Environment with Application: A Case Study of ERP Systems**

Minhaj Afridi, Abdu H. Gumaei , Hussain AlSalman , Asghar Khan , and Sk. Md. Mizanur Rahman







Research Article (23 pages), Article ID 4488576, Volume 2022 (2022)

**[Retracted] A Cost Allocation Decision Model for Air Pollution Control**

Dong-Sheng Qin  and Chang-Yuan Gao 

Research Article (13 pages), Article ID 4116527, Volume 2022 (2022)

**[Retracted] A Double-Layered Belief Rule Base Model for Driving Anger Detection Using Human, Vehicle, and Environment Characteristics: A Naturalistic Experimental Study**

Ping Wan , Xinyan Deng , Lixin Yan , Xiaowei Jing , Liquan Peng , and Xu Wang 


Research Article (17 pages), Article ID 5698393, Volume 2022 (2022)

**Evaluation of Regional Ecological Efficiency and Intelligent Decision Support for Sustainable Development Based on Environmental Big Data**

Lelai Shi , Jing Zhou , and Xingchi Zhou



Research Article (12 pages), Article ID 2820426, Volume 2022 (2022)

**A Big Data-Driven Approach to Analyze the Influencing Factors of Enterprise's Technological Innovation**

Qianqian Zhang 


Research Article (14 pages), Article ID 3785685, Volume 2022 (2022)

**[Retracted] An Individual Driving Behavior Portrait Approach for Professional Driver of HDVs with Naturalistic Driving Data**

Yi He , Shuo Yang, Xiao Zhou , and Xiao-Yun Lu

Research Article (14 pages), Article ID 3970571, Volume 2022 (2022)

**[Retracted] A DSGE Decision Model for Investigating the LPR Transmission Effect**

Xing Yang 

Research Article (10 pages), Article ID 2981558, Volume 2022 (2022)

**Edge Caching in Fog-Based Sensor Networks through Deep Learning-Associated Quantum Computing Framework**

Tayyabah Hasan, Fahad Ahmad , Muhammad Rizwan , Nasser Alshammari , Saad Awadh Alanazi , Iftikhar Hussain, and Shahid Naseem



Research Article (17 pages), Article ID 6138434, Volume 2022 (2022)

**[Retracted] Sliding Mode Control of Flexible Articulated Manipulator Based on Robust Observer**

Yanghua Gao  and Hailiang Lu 




Research Article (10 pages), Article ID 2440770, Volume 2022 (2022)

**[Retracted] Computational Prediction of Intrinsically Disordered Proteins Based on Protein Sequences and Convolutional Neural Networks**

Hao He  and Yong Yang 

Research Article (8 pages), Article ID 4455604, Volume 2021 (2021)

**Modeling the Use of Online Knowledge Community: A Perspective of Needs-Affordances-Features**

Zeyu Jiao , Jianbin Chen , and Eunjin Kim 

Research Article (16 pages), Article ID 3496807, Volume 2021 (2021)

**[Retracted] Cluster-Based Mutual Fund Classification and Price Prediction Using Machine Learning for Robo-Advisors**

Xiaofei Chen, Shujun Ye, and Chao Huang 



Research Article (14 pages), Article ID 4984265, Volume 2021 (2021)

**[Retracted] Generalized Zero-Adjusted Models to Predict Medical Expenditures**

Xin Xu , Tao Ye , and Dongxiao Chu 

Research Article (18 pages), Article ID 5874275, Volume 2021 (2021)

**Is It Possible to Earn Abnormal Return in an Inefficient Market? An Approach Based on Machine Learning in Stock Trading**

Bui Thanh Khoa  and Tran Trong Huynh 


Research Article (14 pages), Article ID 2917577, Volume 2021 (2021)

**[Retracted] A Game Model for Incentive Mechanism of Distributed Nodes in Supply Chains**

Jingjing Jiang  and Aobo Lyu 


Research Article (10 pages), Article ID 7502860, Volume 2021 (2021)

**[Retracted] Modeling on Social Health Literacy Level Prediction**

Xuemei You , Yongdong Liu, Mingming Zhang, Man Zhang, Yangli Yu, and Chengge Sang



Research Article (9 pages), Article ID 7049997, Volume 2021 (2021)

**A Mixed Decision Strategy for Freight and Passenger Transportation in Metro Systems**

Yutao Ye, Junhua Guo , and Lixin Yan

Research Article (22 pages), Article ID 5412016, Volume 2021 (2021)

**[Retracted] Supply Chain Decision Analysis of Community E-Commerce Platform under Different Power Structures: Considering the Influence of Value Cocreation**

Ziyu Liu  and Yaping Li 

Research Article (10 pages), Article ID 2522245, Volume 2021 (2021)

## Contents


---

**[Retracted] A Novel Hybrid Algorithm for Multiobjective Location-Allocation Problem in Emergency Logistics**

Hongrui Chu  and Yahong Chen

Research Article (12 pages), Article ID 1951161, Volume 2021 (2021)

**[Retracted] Intelligent Control of Agricultural Irrigation through Water Demand Prediction Based on Artificial Neural Network**

Qiuyu Bo and Wuqun Cheng 




Research Article (10 pages), Article ID 7414949, Volume 2021 (2021)

**[Retracted] A Data-Driven Control Strategy for Urban Express Ramp**

Liyang Zhang , Min Zhang , Jian Ma , and Jing Ge 


Research Article (12 pages), Article ID 6540972, Volume 2021 (2021)

**On the Generalized Decreasing Mean Time to Failure or Replaced Ordering**

Haiyan Wang , Diantong Kang , and Lei Yan 

Research Article (23 pages), Article ID 7508812, Volume 2021 (2021)

**[Retracted] Security and Privacy Risk Assessment of Energy Big Data in Cloud Environment**

Zhiru Li, Wei Xu, Huibin Shi, Yuanyuan Zhang , and Yan Yan

Research Article (11 pages), Article ID 2398460, Volume 2021 (2021)



## Retraction

# Retracted: A Data-Driven Control Strategy for Urban Express Ramp

### Computational Intelligence and Neuroscience

Received 31 October 2023; Accepted 31 October 2023; Published 1 November 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] L. Zhang, M. Zhang, J. Ma, and J. Ge, "A Data-Driven Control Strategy for Urban Express Ramp," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 6540972, 12 pages, 2021.

## *Retraction*

# **Retracted: Modeling on Social Health Literacy Level Prediction**

### **Computational Intelligence and Neuroscience**

Received 17 October 2023; Accepted 17 October 2023; Published 18 October 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### **References**

- [1] X. You, Y. Liu, M. Zhang, M. Zhang, Y. Yu, and C. Sang, "Modeling on Social Health Literacy Level Prediction," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 7049997, 9 pages, 2021.

## Retraction

# Retracted: Security and Privacy Risk Assessment of Energy Big Data in Cloud Environment

### Computational Intelligence and Neuroscience

Received 17 October 2023; Accepted 17 October 2023; Published 18 October 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] Z. Li, W. Xu, H. Shi, Y. Zhang, and Y. Yan, "Security and Privacy Risk Assessment of Energy Big Data in Cloud Environment," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 2398460, 11 pages, 2021.

## Retraction

# Retracted: Intelligent Control of Agricultural Irrigation through Water Demand Prediction Based on Artificial Neural Network

### Computational Intelligence and Neuroscience

Received 17 October 2023; Accepted 17 October 2023; Published 18 October 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] Q. Bo and W. Cheng, "Intelligent Control of Agricultural Irrigation through Water Demand Prediction Based on Artificial Neural Network," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 7414949, 10 pages, 2021.

## Retraction

# Retracted: A DSGE Decision Model for Investigating the LPR Transmission Effect

### Computational Intelligence and Neuroscience

Received 17 October 2023; Accepted 17 October 2023; Published 18 October 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] X. Yang, "A DSGE Decision Model for Investigating the LPR Transmission Effect," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 2981558, 10 pages, 2022.

## *Retraction*

# **Retracted: Optimal Defect Detection and Sensing System of Railway Tunnel Radar considering Multisensor System Combined with Active Interference Suppression Algorithm**

### **Computational Intelligence and Neuroscience**

Received 17 October 2023; Accepted 17 October 2023; Published 18 October 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### **References**

- [1] Y. Lei, Y. Zou, B. Jiang, and T. Tian, "Optimal Defect Detection and Sensing System of Railway Tunnel Radar considering Multisensor System Combined with Active Interference Suppression Algorithm," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 2459996, 13 pages, 2022.

## *Retraction*

# **Retracted: Development of a Relative Similarity Degree Based Engineering Construction Multi-Attribute Decision Model and Its Application**

### **Computational Intelligence and Neuroscience**

Received 17 October 2023; Accepted 17 October 2023; Published 18 October 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### **References**

- [1] L. Sui, X. Liang, and Q. Huang, "Development of a Relative Similarity Degree Based Engineering Construction Multi-Attribute Decision Model and Its Application," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4274795, 5 pages, 2022.

## *Retraction*

# **Retracted: Cloud Computing Load Balancing Mechanism Taking into Account Load Balancing Ant Colony Optimization Algorithm**

### **Computational Intelligence and Neuroscience**

Received 17 October 2023; Accepted 17 October 2023; Published 18 October 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### **References**

- [1] J. He, "Cloud Computing Load Balancing Mechanism Taking into Account Load Balancing Ant Colony Optimization Algorithm," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3120883, 10 pages, 2022.



## *Retraction*

# **Retracted: A Decision Method for Benefit Distribution Mechanism of Shared End Distribution on Shapley Value**

### **Computational Intelligence and Neuroscience**

Received 17 October 2023; Accepted 17 October 2023; Published 18 October 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### **References**

- [1] H. Zhang, Y. Hou, and F. He, "A Decision Method for Benefit Distribution Mechanism of Shared End Distribution on Shapley Value," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3816440, 13 pages, 2022.

## Retraction

# Retracted: Supply Chain Decision Analysis of Community E-Commerce Platform under Different Power Structures: Considering the Influence of Value Cocreation

### Computational Intelligence and Neuroscience

Received 17 October 2023; Accepted 17 October 2023; Published 18 October 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] Z. Liu and Y. Li, "Supply Chain Decision Analysis of Community E-Commerce Platform under Different Power Structures: Considering the Influence of Value Cocreation," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 2522245, 10 pages, 2021.

## Retraction

# Retracted: Local Privacy Protection for Sensitive Areas in Multiface Images

### Computational Intelligence and Neuroscience

Received 3 October 2023; Accepted 3 October 2023; Published 4 October 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] C. Liu, J. Yang, X. Zhang et al., "Local Privacy Protection for Sensitive Areas in Multiface Images," *Computational Intelligence and Neuroscience*, vol. 2022, 15 pages, 2022.

## Retraction

# Retracted: Generalized Zero-Adjusted Models to Predict Medical Expenditures

### Computational Intelligence and Neuroscience

Received 3 October 2023; Accepted 3 October 2023; Published 4 October 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] X. Xu, T. Ye, and D. Chu, "Generalized Zero-Adjusted Models to Predict Medical Expenditures," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 5874275, 18 pages, 2021.

## Retraction

# Retracted: An Optimized Decision Method for Smart Teaching Effect Based on Cloud Computing and Deep Learning

### Computational Intelligence and Neuroscience

Received 3 October 2023; Accepted 3 October 2023; Published 4 October 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] M. Jiang and Y. Sun, "An Optimized Decision Method for Smart Teaching Effect Based on Cloud Computing and Deep Learning," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 6907172, 10 pages, 2022.

## Retraction

# Retracted: Innovative Research on Urban Community Governance Decision-Making Relying on Distributed High-Performance Computing Blockchain Key Algorithms

### Computational Intelligence and Neuroscience

Received 3 October 2023; Accepted 3 October 2023; Published 4 October 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] J. Wen and P. Xie, "Innovative Research on Urban Community Governance Decision-Making Relying on Distributed High-Performance Computing Blockchain Key Algorithms," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4078014, 13 pages, 2022.

## Retraction

# Retracted: An Individual Driving Behavior Portrait Approach for Professional Driver of HDVs with Naturalistic Driving Data

### Computational Intelligence and Neuroscience

Received 3 October 2023; Accepted 3 October 2023; Published 4 October 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] Y. He, S. Yang, X. Zhou, and X. Lu, "An Individual Driving Behavior Portrait Approach for Professional Driver of HDVs with Naturalistic Driving Data," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3970571, 14 pages, 2022.

## Retraction

# Retracted: A Double-Layered Belief Rule Base Model for Driving Anger Detection Using Human, Vehicle, and Environment Characteristics: A Naturalistic Experimental Study

### Computational Intelligence and Neuroscience

Received 3 October 2023; Accepted 3 October 2023; Published 4 October 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] P. Wan, X. Deng, L. Yan, X. Jing, L. Peng, and X. Wang, "A Double-Layered Belief Rule Base Model for Driving Anger Detection Using Human, Vehicle, and Environment Characteristics: A Naturalistic Experimental Study," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 5698393, 17 pages, 2022.



## Retraction

# Retracted: Sliding Mode Control of Flexible Articulated Manipulator Based on Robust Observer

### Computational Intelligence and Neuroscience

Received 15 August 2023; Accepted 15 August 2023; Published 16 August 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] Y. Gao and H. Lu, "Sliding Mode Control of Flexible Articulated Manipulator Based on Robust Observer," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 2440770, 10 pages, 2022.

## Retraction

# Retracted: Design of Sponsored Search Auction Mechanism for Federated Learning Advertising Platform

### Computational Intelligence and Neuroscience

Received 1 August 2023; Accepted 1 August 2023; Published 2 August 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] H. Jiang, T. Cui, and K. Yang, "Design of Sponsored Search Auction Mechanism for Federated Learning Advertising Platform," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 5787491, 15 pages, 2022.

## Retraction

# Retracted: Tripartite Evolutionary Game Analysis of Financial Support for Science-Tech Enterprises' Innovation with Government Participation

### Computational Intelligence and Neuroscience

Received 1 August 2023; Accepted 1 August 2023; Published 2 August 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] W. Liu, Y. Liu, J. Gu, J. Zhao, and R. Zhang, "Tripartite Evolutionary Game Analysis of Financial Support for Science-Tech Enterprises' Innovation with Government Participation," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 5207003, 21 pages, 2022.

## Retraction

# Retracted: A Study on the Impact of Regional Total Factor Production in Digital Economy Based on Fuzzy Hierarchical VISC Algorithm

### Computational Intelligence and Neuroscience

Received 1 August 2023; Accepted 1 August 2023; Published 2 August 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] H. Li and W. Xue, "A Study on the Impact of Regional Total Factor Production in Digital Economy Based on Fuzzy Hierarchical VISC Algorithm," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 6903836, 11 pages, 2022.

## Retraction

# Retracted: Managing the Gate Assignment Problem in the Hub Airport with Satellite Halls: A Transfer Demand-Oriented Approach

### Computational Intelligence and Neuroscience

Received 1 August 2023; Accepted 1 August 2023; Published 2 August 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] R. Wu, W. Zhou, L. Zhong, and Y. Liu, "Managing the Gate Assignment Problem in the Hub Airport with Satellite Halls: A Transfer Demand-Oriented Approach," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 5087941, 15 pages, 2022.

## Retraction

# Retracted: An Algorithm for Time Prediction Signal Interference Detection Based on the LSTM-SVM Model

### Computational Intelligence and Neuroscience

Received 1 August 2023; Accepted 1 August 2023; Published 2 August 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] N. Xiao and Z. Song, "An Algorithm for Time Prediction Signal Interference Detection Based on the LSTM-SVM Model," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 1626458, 8 pages, 2022.

## Retraction

# Retracted: A Novel Hybrid Algorithm for Multiobjective Location-Allocation Problem in Emergency Logistics

### Computational Intelligence and Neuroscience

Received 1 August 2023; Accepted 1 August 2023; Published 2 August 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] H. Chu and Y. Chen, "A Novel Hybrid Algorithm for Multiobjective Location-Allocation Problem in Emergency Logistics," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 1951161, 12 pages, 2021.

## Retraction

# Retracted: A Prediction Model Analysis of Behavior Recognition Based on Genetic Algorithm and Neural Network

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] Q. Wang and S. Liu, "A Prediction Model Analysis of Behavior Recognition Based on Genetic Algorithm and Neural Network," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3552908, 11 pages, 2022.



## Retraction

# Retracted: Multirobot Collaborative Pursuit Target Robot by Improved MADDPG

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] X. Zhou, S. Zhou, X. Mou, and Y. He, "Multirobot Collaborative Pursuit Target Robot by Improved MADDPG," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4757394, 10 pages, 2022.

## *Retraction*

# **Retracted: Modeling and Optimization of Water Distribution in Mineral Processing considering Water Cost and Recycled Water**

### **Computational Intelligence and Neuroscience**

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### **References**

- [1] H. Wang, Q. Cui, and H. Du, "Modeling and Optimization of Water Distribution in Mineral Processing considering Water Cost and Recycled Water," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 2314788, 12 pages, 2022.

## Retraction

# Retracted: Brand Marketing Leveraging the Advantage of Emoji Pack Relying on Association Rule Algorithm in Data Mining Technology

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] H. Tian, "Brand Marketing Leveraging the Advantage of Emoji Pack Relying on Association Rule Algorithm in Data Mining Technology," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3511211, 9 pages, 2022.

## Retraction

# Retracted: Dynamic Mechanism and Decision Analysis of Urban Exhibition Tourism Development Based on High-Performance Computing

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] H. Wan, "Dynamic Mechanism and Decision Analysis of Urban Exhibition Tourism Development Based on High-Performance Computing," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3806752, 14 pages, 2022.

## Retraction

# Retracted: Picture Fuzzy Einstein Hybrid-Weighted Aggregation Operator and Its Application to Multicriteria Group Decision Making

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] G. Cao, "Picture Fuzzy Einstein Hybrid-Weighted Aggregation Operator and Its Application to Multicriteria Group Decision Making," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 6925670, 20 pages, 2022.

## Retraction

# Retracted: An Annealing Model Analysis and Research on the In of Community Public Sports on Economic Development

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] Z. Guo, "An Annealing Model Analysis and Research on the In of Community Public Sports on Economic Development," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3129638, 9 pages, 2022.

## Retraction

# Retracted: Novel Decision-Making Techniques in Tripolar Fuzzy Environment with Application: A Case Study of ERP Systems

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] M. Afridi, A. H. Gumaei, H. AlSalman, A. Khan, and S. M. Mizanur Rahman, "Novel Decision-Making Techniques in Tripolar Fuzzy Environment with Application: A Case Study of ERP Systems," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4488576, 23 pages, 2022.

## Retraction

# Retracted: Random Forest Algorithm Based on Speech for Early Identification of Parkinson's Disease

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] P. Fan, "Random Forest Algorithm Based on Speech for Early Identification of Parkinson's Disease," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3287068, 6 pages, 2022.



## Retraction

# Retracted: A Cost Allocation Decision Model for Air Pollution Control

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] D. Qin and C. Gao, "A Cost Allocation Decision Model for Air Pollution Control," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4116527, 13 pages, 2022.

## Retraction

# Retracted: Cluster-Based Mutual Fund Classification and Price Prediction Using Machine Learning for Robo-Advisors

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] X. Chen, S. Ye, and C. Huang, "Cluster-Based Mutual Fund Classification and Price Prediction Using Machine Learning for Robo-Advisors," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 4984265, 14 pages, 2021.

## Retraction

# Retracted: Application of Mathematical Models in Economic Variable Input and Output Models under the Scientific Visualization

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] K. Zheng and Y. Liu, "Application of Mathematical Models in Economic Variable Input and Output Models under the Scientific Visualization," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 6269358, 10 pages, 2022.

## Retraction

# Retracted: RFID Data Analysis and Evaluation Based on Big Data and Data Clustering

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] L. Lv, "RFID Data Analysis and Evaluation Based on Big Data and Data Clustering," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3432688, 10 pages, 2022.

## *Retraction*

# **Retracted: Visualization and Analysis of Mapping Knowledge Domain of Heterogeneous Traffic Flow**

### **Computational Intelligence and Neuroscience**

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### **References**

- [1] Y. He, Q. Feng, L. Yan, and X. Lu, "Visualization and Analysis of Mapping Knowledge Domain of Heterogeneous Traffic Flow," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 7754961, 15 pages, 2022.

## Retraction

# Retracted: A Graph Partition-Based Large-Scale Distribution Network Reconfiguration Method

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] Y. Sha, "A Graph Partition-Based Large-Scale Distribution Network Reconfiguration Method," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3169065, 13 pages, 2022.

## Retraction

# Retracted: Research and Design of Docker Technology Based Authority Management System

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] L. You and H. Sun, "Research and Design of Docker Technology Based Authority Management System," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 5325694, 8 pages, 2022.

## Retraction

# Retracted: Differentially Private Singular Value Decomposition for Training Support Vector Machines

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] Z. Sun, J. Yang, and X. Li, "Differentially Private Singular Value Decomposition for Training Support Vector Machines," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 2935975, 11 pages, 2022.



## Retraction

# Retracted: A Mode of Intelligent Equipment Monitoring Optimization Based on Dynamic Programming Algorithm

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] Z. Zheng and C. Jiang, "A Mode of Intelligent Equipment Monitoring Optimization Based on Dynamic Programming Algorithm," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 1569428, 11 pages, 2022.

## *Retraction*

# **Retracted: Digital Protection and Development of Intangible Cultural Heritage Relying on High-Performance Computing**

### **Computational Intelligence and Neuroscience**

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### **References**

- [1] H. Ding, "Digital Protection and Development of Intangible Cultural Heritage Relying on High-Performance Computing," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4955380, 10 pages, 2022.

## Retraction

# Retracted: A Game Model for Incentive Mechanism of Distributed Nodes in Supply Chains

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] J. Jiang and A. Lyu, "A Game Model for Incentive Mechanism of Distributed Nodes in Supply Chains," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 7502860, 10 pages, 2021.

## Retraction

# Retracted: High-Performance Computational Recognition of Communication Signals Based on Bispectral Quadratic Feature Model

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] Y. Chen, R. Zhu, J. Guo, and F. Wang, "High-Performance Computational Recognition of Communication Signals Based on Bispectral Quadratic Feature Model," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 2773492, 12 pages, 2022.

## Retraction

# Retracted: Research on Probability Distribution of Short-Term Photovoltaic Output Forecast Error Based on Numerical Characteristic Clustering

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] P. Yan, C. Xiang, T. Li et al., "Research on Probability Distribution of Short-Term Photovoltaic Output Forecast Error Based on Numerical Characteristic Clustering," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 5355286, 11 pages, 2022.

## Retraction

# Retracted: Computational Prediction of Intrinsically Disordered Proteins Based on Protein Sequences and Convolutional Neural Networks

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] H. He and Y. Yang, "Computational Prediction of Intrinsically Disordered Proteins Based on Protein Sequences and Convolutional Neural Networks," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 4455604, 8 pages, 2021.

## Retraction

# Retracted: Order Management and Completion Date Prediction of Manufacturing Job-Shop Based on Deep Learning

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] M. Wang, "Order Management and Completion Date Prediction of Manufacturing Job-Shop Based on Deep Learning," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3458116, 10 pages, 2022.

## *Retraction*

# **Retracted: Spare Parts Stocking Decision Strategy and Service Logistics Cost Optimization of Two-Echelon Service Logistics System considering Multifailure Mode**

### **Computational Intelligence and Neuroscience**

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### **References**

- [1] R. Wang and G. Chen, "Spare Parts Stocking Decision Strategy and Service Logistics Cost Optimization of Two-Echelon Service Logistics System considering Multifailure Mode," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 7607985, 12 pages, 2022.



## Retraction

# Retracted: A Take-Over Performance Evaluation Model for Automated Vehicles from Automated to Manual Driving

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] L. Yan, J. Chen, C. Wen, P. Wan, L. Peng, and X. Yu, "A Take-Over Performance Evaluation Model for Automated Vehicles from Automated to Manual Driving," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3160449, 16 pages, 2022.

## Retraction

# Retracted: Feature Sequencing Method of Industrial Control Data Set Based on Multidimensional Evaluation Parameters

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] X. Liu, X. Kong, X. Zhang et al., "Feature Sequencing Method of Industrial Control Data Set Based on Multidimensional Evaluation Parameters," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 9248267, 10 pages, 2022.

## Retraction

# Retracted: A Prediction Model of Health Development Based on Linear Sequential Extreme Learning Machine Algorithm Matrix

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] S. Cheng and S. Liu, "A Prediction Model of Health Development Based on Linear Sequential Extreme Learning Machine Algorithm Matrix," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 7632841, 12 pages, 2022.

## Retraction

# Retracted: Discontinuous Track Recognition System Based on PolyLaneNet for Darwin-op2 Robot

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] X. Wu, S. Lv, X. Wang et al., "Discontinuous Track Recognition System Based on PolyLaneNet for Darwin-op2 Robot," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 5431886, 10 pages, 2022.

## Retraction

# Retracted: Dynamic Performance Analysis of STEP System in Internet of Vehicles Based on Queuing Theory

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] H. Li, X. Wu, and Y. Wang, "Dynamic Performance Analysis of STEP System in Internet of Vehicles Based on Queuing Theory," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 8322029, 13 pages, 2022.

## *Retraction*

# **Retracted: Development of a Relative Similarity Degree Based Engineering Construction Multi-Attribute Decision Model and Its Application**

### **Computational Intelligence and Neuroscience**

Received 17 October 2023; Accepted 17 October 2023; Published 18 October 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### **References**

- [1] L. Sui, X. Liang, and Q. Huang, "Development of a Relative Similarity Degree Based Engineering Construction Multi-Attribute Decision Model and Its Application," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4274795, 5 pages, 2022.

## Research Article

# Development of a Relative Similarity Degree Based Engineering Construction Multi-Attribute Decision Model and Its Application

Ling Sui, Xiaojie Liang , and Qilong Huang

No. 501, The Transportation Development Research Center of China Academy of Transportation Sciences, Ministry of Transport R. China, 240 Huixinli, Chaoyang District, Beijing 100029, China

Correspondence should be addressed to Xiaojie Liang; [dmluxj@163.com](mailto:dmluxj@163.com)

Received 5 January 2022; Accepted 18 March 2022; Published 17 June 2022

Academic Editor: Daqing Gong

Copyright © 2022 Ling Sui et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Generally, there are large amounts of uncertain factors in the multi-attribute decision system. By using the gray relational degree and fuzzy gray relational degree, the weights of the comprehensive indexes are extracted. Then, a novel decision model is established based on the concept of relative similarity degree. Finally, comparative research is carried out taking the maritime safety engineering construction in Hunan Province, China, as an example to verify that the developed model is rather effective and practical for its high resolution and sensitivity in multi-attribute decision.

## 1. Introduction

Multi-attribute decision problems refer to selection of the ideal scheme from a limited number of alternative schemes after comprehensively comparing these attributes (indexes) and ranking them in the scheme set. Each attribute in a decision system generally has uncertainty to some extent. Aiming at the presence of these large amounts of uncertain factors, numerous researchers have conducted relevant research. The commonly used method is to extract the weights of indexes to serve the final decision by using the gray relational degree and fuzzy processing method [1, 2]. Literature [3–5] proposed a decision-making method based on possibility and satisfaction, which was improved in literature [6]. Literature [7, 8] gives an integrated method of multi-objective decision-making and extends this method. On this basis, literature [9–15] gives a multi-attribute decision-making method based on distance similarity measure. Literature [16–20] introduces the concept of uncertain fuzzy set and gives an improved method of multiobjective decision-making. Literature [21–23] proposed a new similarity weighted aggregation method.

While these methods can solve practical problems, the obtained decision results show low discreteness and tend to be normalized sometimes, which influences the resolution of

the results and the utilization of given information in the system. The above problem becomes more prominent in the comparison of schemes with similar levels, high fuzzy degree, and difficulty in selection. To solve the problem, the concept of relative similarity degree is introduced to construct a novel decision model, followed by comparative research based on a practical case. The research provides a reasonable and effective approach for the comprehensive evaluation of a project.

## 2. Problem Description

In decision-making problems, in order to ensure the accuracy and effectiveness of decision-making results, the similarity between language terms is a factor that cannot be ignored. In order to better express decision information, the following symbolic notation is used to represent the sets and quantities related to multi-attribute decision-making problems:

$X = \{x_1, x_2, \dots, x_m\}$  represents the set of  $m$  decision schemes

$S = \{s_1, s_2, \dots, s_n\}$  refers to the set of  $n$  attributes which are assumed to be independent



$\mathbf{A} = [a_{ij}]_{m \times n}$  represents the decision matrix, where  $a_{ij}$  is a result of an attribute  $s_j$  using a scheme  $x_i$  (i.e., the attribute value)

In general, attributes can be divided into benefit, cost, fixed, and interval types and different attributes possibly have different dimensions. Therefore, for the sake of convenient analysis and calculation, the decision matrix  $\mathbf{A}$  needs to be normalized, using the normalization formula in Reference [1] for instance. Suppose that the normalized decision matrix is  $\mathbf{B} = [b_{ij}]_{m \times n}$  and  $\bar{b}_i = (1/n) \cdot \sum_{j=1}^n b_{i,j}$  ( $i = 1, \dots, m$ ) is the expected value of the index.

Let  $\omega_k = (\omega_1, \omega_2, \dots, \omega_n)^T$ , where  $\omega_j \geq 0$  ( $j = 1, 2, \dots, n$ ), and  $\sum_{j=1}^n \omega_j = 1$  is the weight vector of an attribute and  $k$  denotes the different methods used for extracting the weight. The concepts of ideal point of multiobjective decision-making problem, negative ideal point of multiobjective decision-making problem, and objective vector corresponding to any feasible solution are introduced. It is extracted by gray correlation degree and fuzzy gray correlation degree operator,  $k = 1, 2$ .

$$r_{i,j} = \frac{\min_{i=1}^m |b_{i,j} - \bar{b}_i| + \zeta \max_{i=1}^m |b_{i,j} - \bar{b}_i|}{|b_{i,j} - \bar{b}_i| + \zeta \max_{i=1}^m |b_{i,j} - \bar{b}_i|} \quad (i = 1, \dots, m; j = 1, \dots, n), \quad (2)$$

is the average gray relational degree. The constant  $\zeta$  is the identification coefficient of gray relation. It is used to adjust the size of the comparative environment and generally let it be  $\zeta = 0.5$  in practical engineering applications.

**3.1.2. Operator of Fuzzy Gray Relation.** According to the theory of fuzzy system, the weight of the  $j$ th index is defined as follows:

$$\omega_j = \frac{\sum_{i=1}^m \tilde{r}_{i,j}}{\sum_{j=1}^n \sum_{i=1}^m \tilde{r}_{i,j}}, \quad (3)$$

where

$$\tilde{r}_{i,j} = \frac{\sum_{i=1}^m b_{i,j} \wedge \bar{b}_i}{\sum_{i=1}^m b_{i,j} \vee \bar{b}_i}. \quad (4)$$

**3.2. Determining the Relative Similarity Degree.** Because the relative closeness makes effective use of the distance information from the structural type scheme to the ideal scheme and the negative ideal scheme, the relative closeness method for multi-attribute decision-making can overcome the limitations brought by only using the Euclidean distance. Using the relative closeness between the objective scheme and the ideal point and the satisfaction of the objective function, an interactive multiobjective decision-making method based on relative closeness is proposed as follows:

- (1) Constructing the expanded normalized decision matrix  $\mathbf{C}$

### 3. Establishment of the Multiattribute Decision Model

**3.1. Extracting Weight  $\omega_j$  of the Index.** Reasonably and correctly extracting the weight of an index is of great significance for investment decision. Theoretically speaking, the larger the influence of an index on other indexes is, the greater the information contained by the index in the system; otherwise, the index has less information. Based on this, the weight of the index is extracted by using the relational operator.

**3.1.1. Operator of Gray Relation.** In accordance with the gray relational theory, the weight of the  $j$ th index is defined as follows:

$$\omega_j = \frac{\sum_{i=1}^m r_{i,j}}{\sum_{j=1}^n \sum_{i=1}^m r_{i,j}}, \quad (1)$$

where

Optimal and worst schemes are the possible ideal and least ideal schemes in a comprehensive evaluation problem.

Assume that the optimal and worst schemes are  $\bar{P} = \{p_1, p_2, \dots, p_n\}$  ( $p_j \geq 0$ ) and  $\bar{Q} = \{q_1, q_2, \dots, q_n\}$  ( $q_j \geq 0$ ), respectively.

After normalization of the index, the normalized decision matrix  $\mathbf{B} = [b_{ij}]_{m \times n}$  is obtained, in which  $0 \leq b_{ij} \leq 1$  and the larger the value of  $b_{ij}$  is, the better a scheme under the current index. Obviously, the optimal scheme in the decision matrix  $\mathbf{B}$  is when the values of all of the normalized indexes are 1, and the worst scheme is when the values of the normalized indexes are 0. While as a matter of fact, these two schemes are difficult to obtain. However, to extract the relative similarity degree, the optimal and worst schemes are also considered, that is, adding two rows (all indexes valuing 1 and 0, separately) in the matrix  $\mathbf{B}$ , thus obtaining the following expanded normalized decision matrix  $\mathbf{C}$ :

$$\mathbf{C} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ c_{m1} & c_{m2} & \cdots & c_{mn} \\ 1 & 1 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \end{bmatrix}. \quad (5)$$



- (2) Weights are assigned to the expanded normalized decision matrix  $C$ , to attain the following weighted expanded normalized matrix  $D$ :

$$D = \begin{matrix} x_1 \\ x_2 \\ \dots \\ x_m \\ \bar{P} \\ \bar{Q} \end{matrix} \begin{bmatrix} d_{11} & d_{12} & \dots & d_{1n} \\ d_{21} & d_{22} & \dots & d_{2n} \\ \dots & \dots & \dots & \dots \\ d_{m1} & d_{m2} & \dots & d_{mn} \\ \omega_1 & \omega_2 & \dots & \omega_n \\ 0 & 0 & \dots & 0 \end{bmatrix} \quad (6)$$

$$= \begin{bmatrix} \omega_1 c_{11} & \omega_2 c_{12} & \dots & \omega_n c_{1n} \\ \omega_1 c_{21} & \omega_2 c_{22} & \dots & \omega_n c_{2n} \\ \dots & \dots & \dots & \dots \\ \omega_1 c_{m1} & \omega_2 c_{m2} & \dots & \omega_n c_{mn} \\ \omega_1 & \omega_2 & \dots & \omega_n \\ 0 & 0 & \dots & 0 \end{bmatrix}$$

- (3) Determining the ideal point  $E^+$  and the negative ideal point  $E^-$

$$\begin{aligned} E^+ &= \max_i \{d_{ij} | j = 1, 2, \dots, n\} \\ &= (E_1^+, E_2^+, \dots, E_n^+), \\ E^- &= \max_j \{d_{ij} | j = 1, 2, \dots, n\} \\ &= (E_1^-, E_2^-, \dots, E_n^-). \end{aligned} \quad (7)$$

It can be seen from the matrix  $D$  that  $E^+ = (\omega_1, \omega_2, \dots, \omega_n)$  and  $E^- = (0, 0, \dots, 0)$ .

- (4) Calculating the relative similarity degree of each scheme with the ideal point:

$$\frac{(E^+ - D_i)(E^+ - E^-)^T}{\|E^+ - E^-\|^2} = 1 - \frac{\sum_j \omega_j d_{ij}}{\sum_j \omega_j^2}, \quad i = 1, 2, \dots, m, \quad (8)$$

where  $D_i = (d_{i1}, d_{i2}, \dots, d_{in})$ .

The value of  $T_i$  reflects the similarity degree of a scheme with the ideal scheme and the disparity with the worst scheme, that is, the relative similarity degree with the ideal point. Apparently,  $T \in [0, 1]$ . The smaller the value of  $T_i$  is, the more similar a decision scheme with the ideal point and, correspondingly, the better the scheme; otherwise, the scheme is worst.

**3.3. Establishing the Decision Model.** In general conditions, the decision model for multi-attribute problems based on gray relation and fuzzy gray relation is as follows:

$$Y_i = \sum_j \omega_j \circ b_{ij}, \quad i = 1, 2, \dots, m, \quad (9)$$

where  $\circ$  represents a certain operation and  $Y_i$  is the array output from the decision results. The larger the value of  $Y_i$  is, the better the scheme.

According to the above analysis, the relative similarity degree based decision model is constructed as follows:

$$Y_i = 1 - \frac{\sum_j \omega_j d_{ij}}{\sum_j \omega_j^2}, \quad i = 1, 2, \dots, m, \quad (10)$$

where  $Y_i$  represents the output array of the decision results. The scheme with a smaller value of  $Y_i$  is more similar to the ideal point; that is, the scheme is better.

#### 4. Case Study

To verify the model established in the research, several decision models for the construction and planning of maritime safety engineering in Hunan Province, China, were compared based on the self-designed investment decision support system of maritime safety engineering construction. Taking the construction planning of Maritime Safety Engineering in Hunan Province as an example, the decision-making models of gray correlation degree, fuzzy gray correlation degree, and relative closeness degree are used for comparative research.

From the perspective of the optimal comprehensive benefit of the construction project of local maritime safety engineering, the index system (the RAW data were derived from the *Construction and Planning Report for the Maritime Safety Engineering in Cities of Hunan Province, China* (November, 1999)) for comprehensive evaluation is constructed (Table 1). The 17 technical and economic indexes (attributes) in the table involve multiple aspects including the technology, economy, environment, and society. By conducting a preliminary feasibility study and expert survey, most of these attributes are benefit-type except for the investment volume per unit length of river bank and the buildings demolished for building per unit length of river bank which are cost-type attributes.

At first, the range transformation method is used to normalize the indexes. Then, the weights  $\omega_1$  and  $\omega_2$  of the index based on gray relational degree and fuzzy gray relational degree are obtained by using (1) and (3), respectively [4].

$$\omega_1 = (0.0593 \ 0.0523 \ 0.0583 \ 0.0649 \ 0.0626 \ 0.0611 \ 0.0701 \ 0.0685 \ 0.0615 \ 0.0555 \ 0.0525 \ 0.0474 \ 0.0594 \ 0.0576 \ 0.0556 \ 0.0583 \ 0.055)^T$$

$$\omega_2 = (0.0531 \ 0.0342 \ 0.0514 \ 0.0794 \ 0.0618 \ 0.0694 \ 0.0862 \ 0.0805 \ 0.0744 \ 0.0579 \ 0.0614 \ 0.0175 \ 0.0546 \ 0.0569 \ 0.0423 \ 0.063 \ 0.056)^T$$

The decision results in Table 2 are obtained by using (9).

The decision results in Table 3 are acquired by using (10).

It can be seen from the above case study that the results obtained using different decision models are basically consistent; that is, Changsha applies the optimal scheme while Zhangjiajie uses the worst one. Using the decision-making methods of gray correlation degree and fuzzy gray correlation degree, a decision-making model of correlation degree is constructed based on the concepts of interval gray and fuzzy gray, but the correlation degree with the ideal scheme is not considered, so the scheme ranking may be

TABLE 1: Technical and economic indexes for the construction and planning of maritime safety engineering in cities of Hunan province.

City	Internal rate of return (%)	Net present value ( $1 \times 10^8$ yuan)	Benefit-cost ratio	Affected population (10,000 people)	Area of flooded cultivated land (10,000 hectares)	Direct economic loss ( $1 \times 10^8$ yuan)	Urban Population (10,000 people)	Urban construction area ( $\text{km}^2$ )	Gross fixed assets ( $1 \times 10^8$ yuan)
Changsha	16.2	6.5018	1.47	51.2	0.82	83.84	132.43	115	303.6
Changde	18	5.9223	1.63	55.45	1.65	106	55.45	35.3	144
Zhangjiajie	14.43	0.4894	1.26	3.888	0.0000625	0.2079	12.96	12.5	26
Yiyang	13.4	0.46	1.14	22.4	0.032	13.4	45.9	29.5	106
Yueyang	14.0	1.073	1.21	22	0.3887	18.544	65.17	55	162.8
Zhuzhou	15.3	1.1818	1.34	24.522	0.1841	87.84	71.19	58	277.3
Hengyang	14.5	1.2183	1.26	19.5	0.4373	18.6	54.17	46	234
Xiangtan	23.6	7.1305	2.31	32.6	0.526	50.4	65	40	122

City	GDP per capita (yuan)	Growth factor of maritime safety criterion	Growth factor of conservation criterion	Population in the planning area protected by unit project quantity (10,000 people)	Planning area protected by unit project quantity ( $\text{km}^2$ )	Investment volume for unit project quantity (10,000 yuan)	Buildings demolished for building unit project quantity (10,000 $\text{m}^2$ )	Environmental benefit coefficient
Changsha	39354.84	10.43	0.4286	0.6217	0.5238	651.4353	0.2071	9
Changde	13201.08	10.76	0.3333	0.5454	1.7935	889.7941	0.2293	9
Zhangjiajie	36010.80	9	9	0.6906	0.8978	659.9275	0.2869	5
Yiyang	18540.31	10.11	0.6667	1.4061	2.4332	779.9513	0.1572	5
Yueyang	31497.24	2.33	0.3333	3.8601	4.1780	1463.3565	0.1558	8
Zhuzhou	27081.05	4.71	0.5385	1.7704	1.7499	840.8024	0.2734	6
Hengyang	15320.69	4.71	0.8182	1.0196	0.8413	683.3180	0.0985	7
Xiangtan	14300	4.71	1.5	0.9942	0.9942	671.2281	0.1146	7

TABLE 2: Decision results obtained using the two different methods.

	Changsha (%)	Changde (%)	Zhangjiajie (%)	Yiyang (%)	Yueyang (%)	Zhuzhou (%)	Hengyang (%)	Xiangtan (%)
$\omega_1$	61.56	49.15	21.70	21.47	41.96	39.80	22.30	39.80
$\omega_2$	67.03	53.03	19.90	23.63	42.70	43.37	24.46	39.17

TABLE 3: Decision results based on the relative similarity degree.

	Changsha (%)	Changde (%)	Zhangjiajie (%)	Yiyang (%)	Yueyang (%)	Zhuzhou (%)	Hengyang (%)	Xiangtan (%)
$\omega_1$	37.03	50.24	80.18	78.54	57.78	59.29	77.01	60.04
$\omega_2$	27.65	44.79	82.34	75.11	57.60	53.64	73.27	61.36

one-sided. The decision-making method of relative closeness degree integrates the relevance closeness degree and can comprehensively consider the relationship between each scheme and positive and negative ideal schemes, so that the decision-making will not deviate from the phenomenon and is more practical. It is not difficult to find from the analysis results and the decision results (Table 2) obtained using the method based on the relative similarity degree are more discrete and have more sensitive ranking, so it is more beneficial to the decision of the scheme, in comparison with the results (Table 3) attained using the other method. It indicates that the decision model based on the relative similarity degree is particularly suitable for situations with schemes of little discrepancies and requiring more elaborate and reliable decision.

## 5. Conclusions

Based on the developed investment decision support system of maritime safety engineering construction, the multi-attribute decision model based on the concept of relative similarity degree is used to carry out numerous case studies. The decision results obtained are objective, which verifies the effectiveness and rationality of the model. Compared with the existing gray relational decision-making model, this decision-making method can comprehensively consider the relationship between each scheme and positive and negative ideal schemes, so that the decision-making will not deviate. With clear concept and reasonable logic and making full use of the decision information of the system, the proposed model is capable of improving the sensitivity and resolution

## Retraction

# Retracted: Brand Marketing Leveraging the Advantage of Emoji Pack Relying on Association Rule Algorithm in Data Mining Technology

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] H. Tian, "Brand Marketing Leveraging the Advantage of Emoji Pack Relying on Association Rule Algorithm in Data Mining Technology," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3511211, 9 pages, 2022.

## Research Article

# Brand Marketing Leveraging the Advantage of Emoji Pack Relying on Association Rule Algorithm in Data Mining Technology

**Huan Tian** 

*Zhuhai College of Science and Technology, Zhuhai, Guangdong, China*

Correspondence should be addressed to Huan Tian; [esthertian082@z cst.edu.cn](mailto:esthertian082@z cst.edu.cn)

Received 27 January 2022; Revised 24 February 2022; Accepted 11 March 2022; Published 24 May 2022

Academic Editor: Daqing Gong

Copyright © 2022 Huan Tian. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The random and disordered configuration of expression package image, text, and context is in line with the emotional consumption trend of young groups, so it has become a kind of image symbol favored by young groups. How to integrate the text and images related to emoticons and carry out emotion analysis in a specific consumer environment is the focus of attention. In view of these needs and defects, this study uses the association rule algorithm in data mining technology to explore brand marketing targeting different consumers. Taking the nonverbal symbols on social media as the research object, this study collects and mines the download ranking and appreciation of the expression package of “WeChat expression open platform.” The persistence of users’ use of expression package has the characteristics of power distribution. When choosing, there are sprouting attributes and dynamic and serialized selection preferences. The advantages of emoticon package are used, brand communication and marketing are strengthened, and effective brand marketing analysis is implemented. The simulation results show that by taking advantage of the strong communication and appeal of the expression package, the brand will help to expand the contact surface of the marketing brand concept. Secondly, as a gathering community, expression packs are divided into different circle groups between and within groups, which helps brand marketing realize precision marketing based on “strong connection” with the help of expression packs and promote the occurrence of consumer behavior.

## 1. Introduction

With the continuous progress of social economy, how to secure more market share in the emerging market is an essential indicator and a goal that many commodities chase after. Various brands have applied different promotional methods to enhance their brand awareness, such as developing different commodities for different age groups of users, engaging different celebrities for endorsement, and innovating various forms or methods of advertising. In fact, these methods have changed the traditional sales model and can improve the sales and develop certain markets to some extent [1, 2]. The development and enrichment of the new media technology have offered a new way of presenting the forms of pictures and texts and a new ideology of communication after reintegration. With regard to the emojis, or Internet expressions, as new text symbols and communication methods, a gradual evolution from colors, drawings,

to expression packs has been realized [3]. However, it should be noted that as the online culture is different from the traditional one, such online expressions may be a fusion between symbols or a fusion of texts and expressions, and this way tends to have the capacity of unlimited expansion in the form of conference codes. Hence, unique personality features can be presented based on specific expressions. On the other hand, emoji packs allow the random splicing and grafting of specific linguistic texts and images to express the specific meaning as needed.

From the aspect of cultural promotion, the emoji pack is a form and culture for the young generation. In a special consumerist contextual environment, the capital has completed penetration at a deeper level, while the value embodiment can be implemented among the youth groups at the consumption and market level according to the specific consumer groups, which ensures that this approach and means are direct, effective, and fast. As a specific visual

content, emoji packs can implement a specific youth-oriented cultural form through the completion of specific collocation of images and texts. Thus, it has become a symbol of communication and is highly popular in the youth network collective. The application of emojis in the brand marketing of commodities is in line with the features of the specific social media era, and it is also a fun way of marketing that can deeply engage the youth [4].

In the process of emoji marketing, a large amount of data are often generated, and it is extremely crucial to carry out effective mining of the information from these data. Mining and extracting the implied knowledge and information from a large amount of disordered data require effective methods and approaches of data mining [5]. Typical methods include data analysis, fusion, and knowledge discovery. The so-called data mining is not targeted at specific databases or simple query and retrieval, but at disordered data. No matter the data are at the micro-, meso-, or macro-scale, the relevant data knowledge can be extracted based on these methods and means [6].

For the purpose of addressing these deficiencies and demands, the business logic of brand marketing by means of emoji packs is first sorted out in this study based on the association rule algorithm in the data mining technology [7]. Secondly, brand recognition and consensus are expanded to promote and carry out marketing of the cosmetics and bath products for different consumer groups. Finally, the marketing word of mouth at different circles and levels is analyzed to implement effective brand marketing analysis by leveraging the advantages of emoji packs, aiming to improve the market share of the corresponding products [8, 9].

## 2. Methodology and Algorithm

In terms of the association rule, it is one of the technical directions of data mining, and its essence is to implement the correlation analysis of the relevance of the data sets by carrying out an effective analysis of the record set based on the given data set. This approach has important and extensive applications in finance, justice, business, and other fields.

In the process of using the specific association rule, the corresponding rule can be deemed in the form of “how the condition is considered to be is how the result will turn out.”

The specific association rule is represented by “A  $\rightarrow$  B,” which can be divided into two specific parts: A on the left side is the conditional part and B on the backside is the result part. The conditional part can contain one or more conditions. Under this premise, if the result is true, then the previous condition must be true.

A high coverage rate indicates that the rule is used frequently.

Association analysis can be described mathematically and formally as follows:

- (1) It contains a certain degree of support; that is, at least a certain number of data sets are included in the database

- (2) It contains a certain degree of confidence; that is, while  $X$  records are included in the database, the data set  $Y$  is also included

*2.1. Related Algorithms and Basic Principles.* The association rule is one of the main contents of the data mining algorithms. Based on the association rule, meaningful association information can be identified in a large amount of data; that is, information mining can be carried out based on the frequency of occurrence between the sets of items.

It is assumed that data set  $A$  contains  $n$  samples and  $L$  stands for the set of labels corresponding to the data set. Each sample in  $A$  is represented by a set of feature values (feature) and a label (label). Let  $F$  be the set of feature values in the data set. The specific formula for the category association rule is described as follows:

$$\text{feature} \Rightarrow \text{label}. \quad (1)$$

In the above equation, feature stands for a set of feature values,  $\text{feature} \subseteq F$ , and  $\text{label} \in L$ .

The two commonly used evaluation criteria for the association rules are support and confidence. The specific formula for their calculation is as follows:

$$\text{Support}(F \Rightarrow L) = \frac{|F \cup L|}{|A|}. \quad (2)$$

Support indicates the frequency of occurrence of a particular rule, which determines the frequency of occurrence of the item set  $|F \cup L|$ .

Confidence indicates the degree of trust of the corresponding rule. It is defined as  $\text{Conf}(F \Rightarrow L)$ , and its calculation formula is shown as follows:

$$\text{conf}(F \Rightarrow L) = \frac{\text{support}(F \cup L)}{\text{support}(F)} \times 100\%. \quad (3)$$

The basic structure of data mining for the association rules is shown in Figure 1.

The mining of frequent item sets is an essential but highly time-consuming process in the association analysis process. The FP-growth algorithm is often selected as the association rule analysis algorithm in the industry, which can avoid the generation of a large number of candidate frequent sets, and it is required to implement only two iterations of the database for the copper welding wire to complete the determination of the frequent item set. The FP-growth data mining process can be divided into two steps as follows: (i) construction of the FP-tree and (ii) frequent pattern mining based on the FP-tree, and subsequent mining of the constructed FP-tree to obtain all the frequent patterns. The construction process of FP-tree is described in Figure 2.

It can be observed from the construction process in Figure 2 that the basic conditional pattern is implemented from the most initial frequent items, and the corresponding conditional FP-tree is established on this basis to implement the filtering of frequent items that fail to comply with the corresponding threshold value. In addition, iterative recursive analysis and mining are carried out based on the

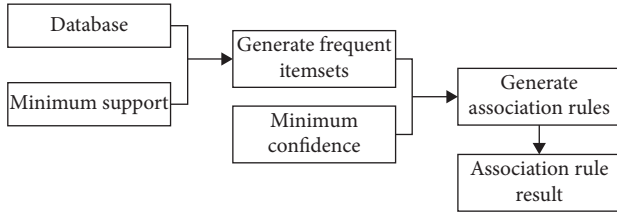


FIGURE 1: Basic structure of the association rules.

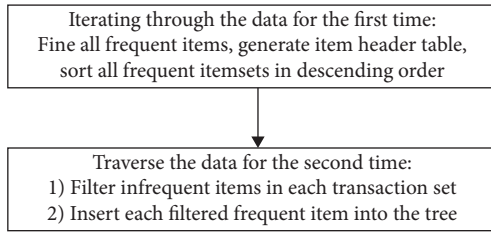


FIGURE 2: FP-tree construction flow.

specific conditional FP-tree until the final FP-tree is void. Subsequently, in the path analysis, the specific frequent items of the frequent item data set are obtained.

The smart grid dispatching technology support system adopts the Global Positioning System (GPS) and has the synchronous timing capability, which has implemented the unified spatial and temporal measurement of large power grids over a wide area. The accumulated wide-area spatial and temporal measurement information in the smart grid dispatching technology support system can be accumulated for a long period (days, months, and years), which can better reflect the spatial and temporal correlation features of the dispatching operation data and therefore have great practical value [10–13].

At present, China has formed a multilevel and hierarchical dispatching operation model at the national, regional, provincial, municipal, and county levels. Based on the scope of dispatch management, responsibilities, geographical location, voltage level, and features of power system, the dispatch functions and priorities at various levels can be different. The national dispatching focuses on UHV AC/DC transmission capacity and emergency guarantee capacity; the regional dispatching attaches more importance to the key contact sections of regional power grids; the provincial dispatching directly controls the output and load of key power plants in the region. At the same time, the security and stability of the power grid have also demonstrated typical regional features and spatial correlation. In this study, a multilevel correlation analysis model is established based on the safety and stability issues at the dispatching level. The details of the model are described in Figure 3. The hierarchical correlation analysis model is established based on empirical and historical data to derive the key factors that may affect the stability of the power grid at various levels and perform selective optimization of the combined input feature space.

Time correlation is the primary feature in the operation of the electric power system, and its correlation is mainly reflected in the following two aspects:

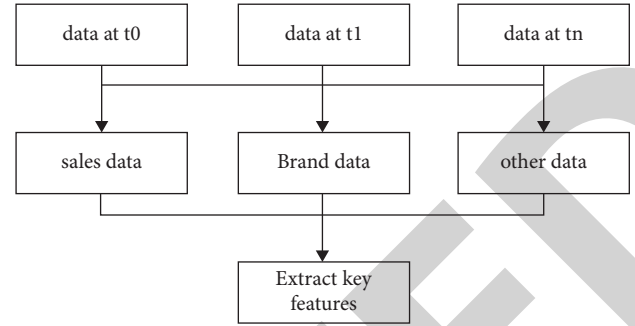


FIGURE 3: Multilevel correlation analysis model.

- (1) The changes in the operation mode of the power grid are closely related to time, and both the output curve and load curve present certain time regularity.
- (2) A large amount of operation data collected by the dispatching center, as well as the calculation data and result data, have a time scale. The data described above are the responses of various dynamic indicators of the power system at various moments, and they all have the corresponding time-series features [14, 15].

It is necessary to integrate the time-series features of the data closely when correlation analysis is carried out. In general, it is highly challenging to analyze the whole time series in data mining. In this study, the operating data are first divided into multiple subsegments according to the time-series features; subsequently, different features are established for each time-series segment, and the association analysis models are constructed separately to obtain the specific association rule models for various time-series segments.

**2.2. Data Mining Technology Association Rule Algorithm.** The essence of the Apriori algorithm is about the application of the database frequent item identification to implement the specific connection and pruning operation extraction, so as to achieve strong regular item set screening that complies with the threshold. The specific calculation formula is as follows:

$$S\% = Sup(X \longrightarrow Y) = P(XY) = \frac{\text{number}(X \cup Y)}{\text{num}(\text{all samples})} \quad (4)$$

The specific calculation of the confidence level is shown as follows:

$$C\% = Conf(X \longrightarrow Y) = P(Y|X) = \frac{P(XY)}{P(X)} \quad (5)$$

If an item set complies with the minimum support, it is referred to as a frequent item set; if it complies with both the minimum support and the minimum confidence, it is referred to as a strong association rule.

In general, the operation of the Apriori algorithm can be divided into 2 stages: (1) the candidate set 1 is generated, and the corresponding support degree is calculated. The unqualified item sets are pruned off according to the set threshold to obtain the frequent item set 1+. (2) The candidate set 2 item sets are generated for item set 1+ connection, and the item sets are further pruned according to the threshold value. Iteration is carried out repeatedly until the frequent item set is valid, and the algorithm is ended. This process is often inefficient due to the necessity to traverse the data set several times. In addition, there are usually a large number of duplicate and invalid misleading rules among the extracted rules. The lift can reflect the relevance of the antecedent and the consequent parts of a rule, while the interest degree indicates the interestingness of a rule [16]. Through the application of the lift, the interest formula can be established. In this way, a large number of invalid rules generated by the Apriori algorithm can be filtered out effectively. The lift is expressed as  $\text{Lift}(X \rightarrow Y)$ , and the interest is expressed as  $\text{Interest}(X \rightarrow Y)$ . Their specific calculation is shown as follows:

$$\text{Lift}(X \rightarrow Y) = \frac{P(X \cup Y)}{P(X)P(Y)}, \quad (6)$$

$$\text{Interest}(X \rightarrow Y) = \frac{\text{Lift}(X \rightarrow Y) - 1}{\text{Lift}(X \rightarrow Y) + 1}. \quad (7)$$

As a class of global optimization-seeking evolutionary algorithms, this feature of genetic algorithms from local to global can largely resolve the problem of inefficiency of the Apriori algorithm. However, it also has some problems. For example, it is easily limited to the local optimal solution, and its efficiency decreases gradually as the algorithm continues.

The ultimate goal of the association rule is to identify the interesting rules that comply with the min-sup and min-conf from a large set of transactions [17, 18]. For the purpose of screening out these interesting rules, the support degree is used to build the fitness function  $(X \rightarrow Y)_1$ ; at the same time, the confidence degree is used to construct the fitness function  $(X \rightarrow Y)_2$  so as to avoid the algorithm from repeated calculations leading to low efficiency, as shown in the following equations:

$$\text{Fitness}(X \rightarrow Y)_1 = \frac{\text{Sup}(X \rightarrow Y)}{\text{min} - \text{sup}}, \quad (8)$$

$$\text{Fitness}(X \rightarrow Y)_2 = \frac{\text{Conf}(X \rightarrow Y)}{\text{min} - \text{conf}}. \quad (9)$$

The selection operation is conducted to ensure the superiority of individuals in the population and to select the individuals with excellent environmental adaptability and pass them to the next generation. In general, the judgment condition of selection is determined according to the fitness function.

The crossover operator and genetic operator in the adaptive genetic algorithm (AGA) are changing in a dynamic manner, which has effectively solved the prematurity problem of the GA. However, the improved operator it has

put forward is highly unfavorable for the iteration of the initial population. On this basis, some scholars have further optimized an adaptive genetic algorithm; the formulas for the calculation of the specific crossover operator  $P_c$  and genetic operator  $P_m$  of this algorithm are shown as follows:

$$P_c = \begin{cases} P_{c1} - \frac{(P_{c1} - P_{c2})(f' - f_{avg})}{f_{max} - f_{avg}}, & f' \geq f_{avg}, \\ P_{c1}, & f' < f_{avg}, \end{cases} \quad (10)$$

$$P_m = \begin{cases} P_{m1} - \frac{(P_{m1} - P_{m2})(f' - f_{avg})}{f_{max} - f_{avg}}, & f' \geq f_{avg}, \\ P_{m1}, & f' < f_{avg}. \end{cases} \quad (11)$$

In this algorithm, the crossover probability and variation probability are improved. In this way, the algorithm will not fall into local solutions at the beginning of iteration. However, in the association rule mining problem, the change in  $(f' - f_{avg})/f_{max} - f_{avg}$  is actually not evident; the dynamic adjustment is not significant. As a result, it is impossible to seek the optimal solution properly. At the beginning of the iteration, it is often expected that the evolution can be accelerated. Hence, relatively high crossover probability and low variation probability are required [19, 20]. As the number of iterations increases, the similarity between individuals of the population will become higher, which can cause the genetic algorithm to fall into the local solution. Thus, it is expected that the genetic algorithm can jump out of the local solution with relatively low crossover probability and high variation probability. In the iterative process, if the crossover probability gradually becomes smaller and the variation probability gradually increases, the problem of premature and inefficient traditional genetic algorithm can be resolved properly. This variation trend is highly similar to the change in the sigmoid function. Hence, this function is combined with the number of iterations to improve the crossover operator  $P_c$  and genetic operator  $P_m$ . The specific calculation formula is as follows:

$$P_c = \begin{cases} P_{c2} + \frac{(P_{c1} - P_{c2})}{1 + e^{G/f_{max}}}, & f' \geq f_{avg}, \\ P_{c1}, & f' < f_{avg}, \end{cases} \quad (12)$$

$$P_m = \begin{cases} P_{m2} + \frac{(P_{m1} - P_{m2})}{1 + e^{G/f_{max}}}, & f' \geq f_{avg}, \\ P_{m1}, & f' < f_{avg}. \end{cases} \quad (13)$$

In the above equation,  $G = t_{max} - t/t_{max}$ ;  $t$  stands for the current number of iterations; and  $t_{max}$  stands for the set of maximum number of iterations.

The detailed steps of the association rule mining algorithm combined with the improved adaptive genetic algorithm and the Apriori algorithm are described as follows.



In Step 1, the database is traversed to generate frequent item sets, which are sorted accordingly. The set of populations is initialized, and each individual in the set of populations with individual bits is encoded, with each real number corresponding to the field of the database, respectively.

In Step 2, the value of each initial population set is calculated according to the corresponding formula.

In Step 3, the selection operation is performed. The specific function calculated value is passed down to the next generation to conduct the corresponding calculation; otherwise, the crossover operation is performed, and the number of specific individuals in the next generation is retained as  $m1$ .

In Step 4, crossover operation is conducted. Two different individuals are randomly selected, and the crossover operation is carried out according to the corresponding formula. The individuals after the crossover are put into the next population, and the number of individuals ( $m2$ ) that have been passed down to the next generation after this step is recorded.

In Step 5, the mutation operation is conducted. On the basis of Step 4, the specific individuals are selected, and the variation operation is conducted using the formula to change the specific number of mutations to implement the change in the next-generation population. As a result, the specific number of individuals in the next generation is expressed as  $m3$ .

In Step 6, statistics are carried out on the number of individuals entering the next generation  $m = m1 + m2 + m3$ . If the statistical result fails to meet the specified population scale,  $(n - m)$  individuals are randomly generated to supplement the next generation; then, Step 2 is proceeded to repeat the operation.

In Step 7, whether the condition for the end of the algorithm is met is determined, and the corresponding rule is output at the same time; otherwise, Step 2 is proceeded to continue the loop.

In Step 8, the association rule that complies with the minimum interest degree is output. The operation flow of the algorithm established on the basis of the above steps is shown in Figure 4.

The transaction database used in the algorithm is stored in the event-store, where the population size is life-count, and the termination condition of the algorithm is iteration-max (i.e., the maximum number of iterations). The current population and new-population are used to store the current population and the next generation of new populations, and the final output value is a strong association rule  $T$  that complies with the minimum interest (min-inte) [21, 22]. Thus, the pseudo-code of the algorithm can be described in Algorithm 1.

**2.3. Brand Marketing Leveraging the Creative Communication Advantages of Two-Dimensional IP.** The marketing of commodity brands by means of emoji packs allows the multilevel exposure of the commodity brands in an all-round way, which can enrich the connotation of the brands, facilitate the effective communication of the brands, and achieve the

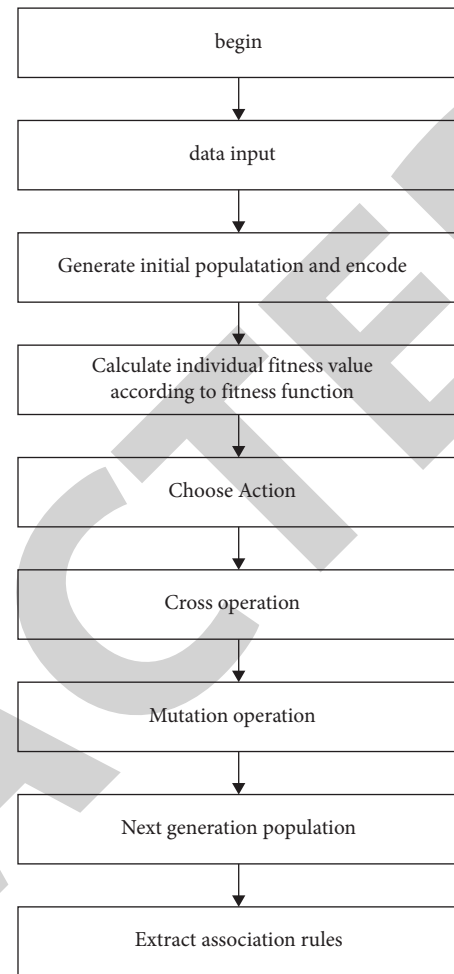


FIGURE 4: Algorithm flow chart.

tension and dimensional expressiveness of advertising creativity [15, 23]. Moreover, emoji packs often come with a certain core audience, which has clarified the specific user groups for the corresponding brands. In addition to the corresponding consumer screening and fixing links, the core value construction and image shaping of commodity brands can build a bridge between the products and consumers.

The rise of various new media application forms has weakened the channel advantages of traditional media; in the context of media decentralization, it is highly challenging to achieve the accurate dissemination of information by mass communication from the perspective of the traditional media. Hence, it is necessary to integrate various types of data to implement the integrated marketing communication of brand information.

Hence, brands can take the advantage of two-dimensional IP in their communication using the cross-media narrative concept. On the basis of traditional integrated marketing communication, advertisers can make use of the IP effect to integrate the discrete marketing mechanisms of each link into an integrated marketing system with a coherence, systemic, and laser-sharp focus.

“The purpose of the communication creative point is to communicate and share, which is the bridge that connects



```

(1) Input: event-store, min-sup, min-conf, life-count, iteration-max
(2) Output: set  $T$  of the strong association rules that comply with the minimum interest threshold (min-inte)
(3) For each  $x$  in event-store do
    //Seek frequent items and carry out sorting, select the individuals that comply with the life-count and store them in the cur-
    population.
    Population = FrequentOne ( $x$ ).
    Cur-population ([0: life-count].
    End for.
(4) For each individual in cur-population do.
    If fitness (individual)  $\geq 1$ .
    New-population (selection (individual).
    Else pass.
    End if.
    End for.
(5) For each two individual (parent1, parent2) (cur-population do
    New-population (new-population  $\cup$  crossover (parent1, parent2).
    //The crossed individuals are removed from the current population.
    For each individual (cur-population do.
    Mutation (individual).
    New-population (new-population  $\cup$  mutation(individual).
    End for.
     $T$  ( $T \cup$  new-population.
    Cur-population (selection (new-population).
(6) End while.
(7) Return  $T$ .

```

ALGORITHM 1: Association Rules Mining Algorithms.

the brand touchpoints and core creative points. Through the integration of hot topics and consumer interests, consumers are ultimately guided to the core brand value. In this process, audience engagement is of crucial significance to the communication of the core value of a brand.”

From the aspect of the coverage of communication, the strong connection based on the community content will form word-of-mouth communication, which in turn will lead to multi-community information dissemination with human communication nodes. The reason is that individuals do not join a single community, but rather participate in the role-playing in a diverse community, which indicates that the individuals prefer to cruise among multiple communities.

The purpose of the core creative points is to create an emotional resonance and value recognition with consumers. As the two-dimensional space culture itself is spontaneously produced by youth groups, it is possible for brands to accomplish the communication of the core creative points smoothly by leveraging the advantages of two-dimensional IP [24, 25].

The value of commodities and services that are endowed with the thoughts of consumers goes beyond the consumption itself and has become a way to acquire self-identity through consumption, which is also the reason why consumers are more willing to pay for such commodities.

### 3. Simulation Experiment

According to the data, 74.19% of Chinese college students understand the application of expression pack in video/short video, and 67.74% of Chinese college students understand

the application of expression pack in comics; 62% of foreign students understand the application of expression pack in comics, and 60% of foreign students understand the application of expression pack in video/short video. However, only 48.39% of Chinese college students and 42% of foreign college students said they had understood the application of expression pack in advertising. It can be seen that the expression package has a high exposure in new media matrices such as short videos and comics, which makes a deep impression on college students. Relatively speaking, the expression package used for advertising scenes also has a certain degree of exposure, but it is relatively weak, indicating that the expression package used for advertising scenes has great market potential and broad market prospects, as shown in Figure 5.

The data mining technology association rule algorithm provides a guarantee for the complete protection of emoji pack, digital restoration and reproduction technology provides support for the effective inheritance of emoji pack, digital display and dissemination technology provides a platform for the widespread sharing of emoji packs, and virtual reality technology provides the space for the development and application of the expression packs. It is protected and implemented to market domestic emoji packs and facilitate the formulation of relevant policies. This study fully analyzes the problems in the inheritance and protection process of emoji pack and applies data mining technology association rule algorithm to the marketing process of emoticons, which provides a strong basis for the realization of emoji pack marketing. The accurate marketing of emoticons can be completed using the data mining technology

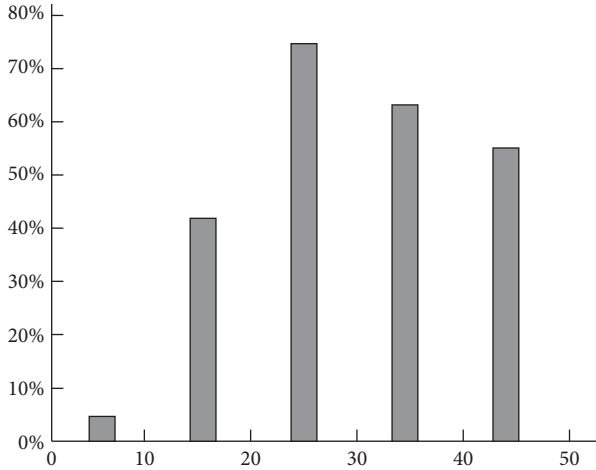


FIGURE 5: College students' understanding of the application of expression pack.

association rule algorithm. In addition, a marketing method of leveraging the advantage of brand emoji pack is also designed.

The data mining technology association rule algorithm is used to realize the marketing of Chinese painting, calligraphy emoji pack, and image package, to train various types of emoji pack suit, and to record the values of  $w_1$ ,  $w_2$ , and  $w_3$  when they reached the maximum value. Figures 6 and 7 show the weight changes, where “o” represents the weight change in smoothness, “—” represents the weight change in consistency, and “\*” represents the weight change in entropy. After training all emoji packs, the average of the two emoji pack values and the average of  $w_1$ ,  $w_2$ , and  $w_3$  are used to test the emoji pack with the average weight value, and the  $I$  value of the test emoji pack is obtained. The data shown in Table 1 (retaining only the integer part) are 10 data arbitrarily extracted from the experimental results. It can be seen from the table that the  $I$  value of the image is relatively small.

For the  $I$  value obtained from the test emoji, the difference measurement of the two emoji packs is carried out as follows:

$$S_i = |I_{\text{test}} - \bar{I}_i|, \quad (14)$$

where  $i = 1, 2$ ,  $\bar{I}_i$  is the average obtained for the two types of emoji packs, and  $I_{\text{test}}$  is the  $I$  value generated by the test emoji pack. In  $S_1 > S_2$ , the test emoji pack is set as the first category; otherwise, it is set as the second category. In  $S_1 = S_2$ , the emoji pack cannot be marketed.

The experimental results are shown in Table 2. The experimental results show that the three groups of experiments have achieved good results, realizing the marketing of Chinese painting, calligraphy emoji pack, and images (stick drawings, tables, function emoji packs). Experiments also fully proved the reliability of the method.

The data mining technology association rule algorithm provides guarantee for the complete protection of expression package, digital restoration and reproduction technology provides support for the effective inheritance of expression package, digital display and communication technology provides a platform for the wide sharing of expression

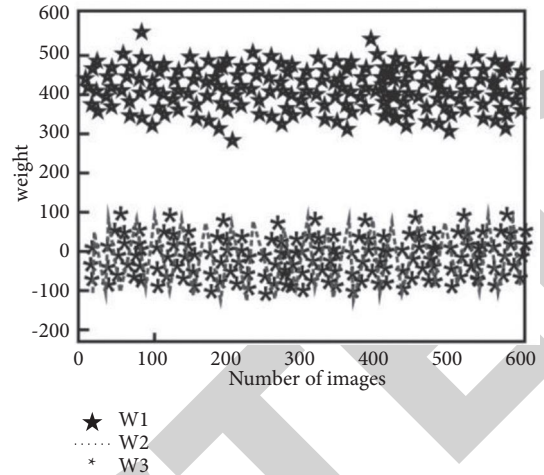


FIGURE 6: Three weight changes of 600 emoji packs.

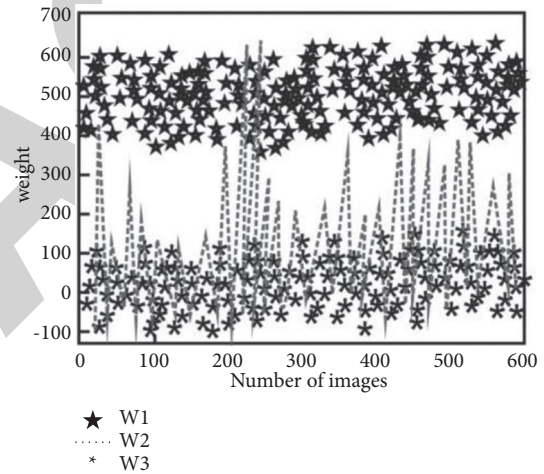


FIGURE 7: Three weight changes of 600 images.

TABLE 1: Comparison of  $I$  values between emoji pack and images.

Emoji pack $I$ value	Image $I$ value	Emoji pack $I$ value	Image $I$ value
3568	1367	3566	767
3656	555	3443	2294
3463	846	-	-

package, and virtual reality technology provides space for the development and utilization of expression package to market the expression package in China and facilitate the formulation, protection, and implementation of relevant policies. This study fully analyzes the problems existing in the inheritance and protection process of expression package and applies the association rule algorithm of data mining technology to the marketing process of expression package, which provides a strong basis for the realization of expression package marketing. Using data mining technology and association rule algorithm, we can complete the accurate marketing of expression package. In addition, we also design the marketing method of using brand expression package.

TABLE 2: Experimental results.

Experiment	Number of emoji packs	Emoji pack		Image		Accuracy rate (%)
		Right	Wrong	Right	Wrong	
1	21	19	0	0	2	90.02
2	15	1	0	13	1	92.87
3	17	8	0	8	1	93.74

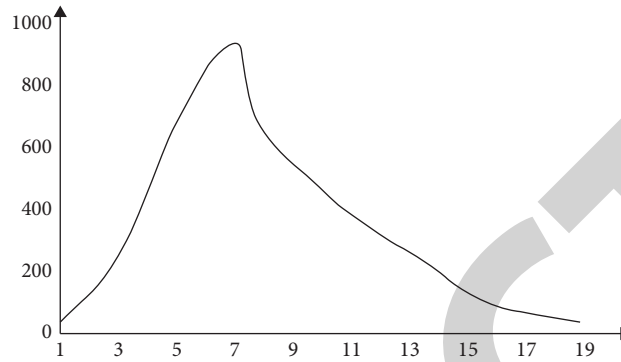


FIGURE 8: Risk frequency chart.

Based on the classification of the specifically defined risk level, the frequency of risks accumulated by the samples over a period of five months is shown in Figure 8.

The sales operation samples on a certain day are divided into five stages according to the time correlation, and the actual sales network data are mined at different voltage levels of the hierarchical data by establishing a model to explore its patterns separately [26]. At the same time, taking into consideration the requirement to extract as much hazardous data as possible, the minimum confidence threshold is set to 35%. In the practical operation of the sales network, the occurrence frequency of severe overload cases is relatively low. Hence, the minimum support threshold of the model is set to 0.9%. Among them, rule 1 is applicable to the first stage; rule 2 is applicable to the second stage; and rule 3 is applicable to the third stage.

From the analysis above, it can be known that the algorithm constructed in this study is in line with the brand marketing by leveraging the advantages of emoji pack. In addition, it can also assist in the identification of some unknown and potential operation patterns in the sales network at the same time and provide a tool for decision-making in the future work of sales operations.

#### 4. Conclusions

By taking advantage of the strong communication and appeal of the expression package, the brand helps to expand the contact surface of the marketing brand concept. In this study, the data mining technology association rule algorithm excavates the association information with high confidence and support, reflects the complexity and interest of the data in the database, and then excavates the beneficial association between the expression package data, promotes the development of the expression package database information mining technology, and fully analyzes the impact of the advantages of the expression package on brand marketing.

The marketing hotspots of the brand for different users are expanded to enhance the reputation and identification points of the brand, and the brand marketing is effectively analyzed using the advantages of emoticon package. The simulation results show that the research on brand expression package under the association rule algorithm of data mining technology is in line with the marketing concept of taking consumers as the core. It can not only help to expand the brand contact points and strengthen the communication creative points using the word-of-mouth marketing of circle communication but also use the emotional consumption characteristics of expression package to stimulate the emotional resonance of the audience and realize the transmission of the core value of the brand. The advantages of emoticon package are fully used to provide support for the application of brand marketing.

#### Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

#### Conflicts of Interest

The authors declare no conflicts of interest.

#### References

- [1] T. De Whirst, "Into the black: marlboro brand architecture, packaging and marketing communication of relative harm," *Tobacco Control*, vol. 27, no. 2, pp. 240–242, 2018.
- [2] A. E. Epperson, L. Henriksen, and J. J. Prochaska, "Natural American spirit brand marketing casts health halo around smoking," *American Journal of Public Health*, vol. 107, no. 3, pp. 19–28, 2017.
- [3] A. J. Moran, E. B. Rimm, and E. M. Taveras, "A school-based brand marketing program's adherence to federal nutrition criteria," *American Journal of Preventive Medicine*, vol. 17, no. 2, pp. 101–108, 2017.

## Research Article

# A Novel U-Net Based Deep Learning Method for 3D Cardiovascular MRI Segmentation

Yinan Lu <sup>1</sup>, Yan Zhao <sup>1</sup>, Xing Chen <sup>2</sup> and Xiaoxin Guo <sup>1</sup>

<sup>1</sup>College of Computer Science and Technology, Jilin University, Changchun 130000, China

<sup>2</sup>College of Artificial Intelligence, Jilin University, Changchun 130000, China

Correspondence should be addressed to Yinan Lu; [luyn@jlu.edu.cn](mailto:luyn@jlu.edu.cn)

Received 9 December 2021; Revised 25 January 2022; Accepted 16 February 2022; Published 20 May 2022

Academic Editor: Daqing Gong

Copyright © 2022 Yinan Lu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Medical multiobjective image segmentation aims to group pixels to form multiple regions based on the different properties of the medical images. Segmenting the 3D cardiovascular magnetic resonance (CMR) images is still a challenging task owing to several reasons, including individual differences in heart shapes, varying signal intensities, and differences in data signal-to-noise ratios. This paper proposes a novel and efficient U-Net-based 3D sparse convolutional network named SparseVoxNet. In this network, there are direct connections between any two layers with the same feature-map size, and the number of connections is reduced. Therefore, the SparseVoxNet can effectively cope with the optimization problem of gradients vanishing when training a 3D deep neural network model on small sample data by significantly decreasing the network depth, and achieving better feature representation using a spatial self-attention mechanism finally. The proposed method in this paper has been thoroughly evaluated on the HVSMR 2016 dataset. Compared with other methods, the method achieves better performance.

## 1. Introduction

In multiobjective segmentation, medical image segmentation aims to segment the images into multiple regions and extract the parts of interest based on the similar characteristics or single attributes of the image, such as edge contour, structure, and shape, which is of great significance for medical image analysis, disease diagnosis, and clinical applications (e.g., 3D computed tomography (CT) and magnetic resonance image (MRI)). Accurate segmentation can not only help precise diagnosis and prediction of prognosis but also benefit surgical planning and intraoperative guidance.

For example, in diagnosing congenital heart disease, segmenting the blood pool and myocardium from 3D cardiovascular magnetic resonance (CMR) images is a prerequisite before creating patient-specific heart models for preprocedural planning of children with complex congenital heart disease (CHD).

One of the main applications of deep learning is medical field, including biomedicine and MRI analysis [1]. However, the level of doctors is uneven, and some departments are

labor-intensive, which has had a profound impact on the development of artificial intelligence in this field [2]. Currently, segmenting vital organs or structures from 3D medical images is an imperative preliminary action for a wide range of clinical treatments. The recognized standard segmentation results are obtained from experienced physicians and radiologists via visual inspection and manual delineations. On the one hand, there always are hundreds of images in an individual's cardiac MRI. It is tedious, time-consuming, and costly to annotate the 3D medical images in a slice-by-slice manner. On the other hand, the whole heart's manual labeling is subjective and suffers from low reproducibility. The results of the labeling could be seriously affected by the experience and knowledge of the observer. Consequently, automatic medical image segmentation with high accuracy is highly demanded. However, using deep learning for automatic medical image segmentation with high accuracy is also a huge challenge [3]. The reasons include (I) the missing borders or indefinite boundaries, with inadequate edge information, and (II) the too low quality of the cardiac images.

There are many deep learning models applied to image segmentation [4]. Convolutional neural network (CNN) based deep learning strategies especially achieved remarkable success in medical image segmentation methods. U-Net [5] is a semantic segmentation network based on fully convolutional networks (FCN) [6], which mainly applied CNN structure for the heart segmentation. U-Net can be quickly trained on small sample data in medical segmentation by data augmentation and achieve outstanding segmentation results. Different from the FCN structure, there is not any encoder or decoder in the U-Net. U-Net contains two paths, the downsampling contraction path, which extracts the high-level abstract features of pixels, and the extended upsampling path, which can reconstruct pixel information lost during downsampling. In the process adopted above, the parts from the comprehensive upsampling approach and the features extracted by downsampling are stitched to maximize the retention of low-level feature information lost by the pooling and convolution operations. Compared with FCN, U-Net can run more efficiently because there is no fully connected layer in the structure. The paper on DenseNet [7] was voted the best paper of CVPR in 2017, which has the same basic idea as ResNet [8]. However, it establishes the dense connection between all the previous layers and the latter layers. The dense block in the DenseNet is a densely connected network model between layers. In each dense block, the input of each layer is the union of the outputs of all the previous layers. DenseNet enhances feature representation with skip connections. However, the feature maps in DenseNet are relatively large, resulting in a large amount of computation in the convolution process, which affects the overall performance of the network.

This paper proposes a novel and efficient 3D sparse convolutional network named SparseVoxNet to comprehensively address these challenges, which can effectively carry out voxel-to-voxel learning and infer 3D medical images. Specifically, we develop a sparse convolutional network that aims to contribute the following ideas:

(1) The sparse network can eliminate redundant computation, reduce model parameters, and decrease the risk of overfitting small sample training data. (2) The full skip connection mechanism in the module can effectively solve the problem of gradient disappearance in 3D deep model training, accelerate the convergence speed, and improve recognition ability. (3) The self-attention mechanism is added to optimize the expression ability of feature maps and capture the long-range dependency between features better.

## 2. Related Work

Multiobjective image segmentation can be divided into supervised and unsupervised methods. Pham et al. [9] proposed a multiobjective optimization approach to segment the brain MRI using fuzzy entropy clustering and region-based active contour methods. Hongwei et al. [10] proposed a multiobjective clustering and toroidal model-guided tracking method to distinguish vascular structures from complicated structures in background regions. In

recent years, deep learning has been successfully applied to medical image segmentation. Çiçek et al. [11] proposed a 3D U-Net network structure to realize the 3D image segmentation. Habijan et al. [12] proposed a framework consisting of two 3D U-Nets. In this framework, the first network was used for localizing the bounding box encompassing the heart, and the second network was employed to segment the different substructures. Ding et al. [13] incorporated attention mechanism within the gradient expanding process to enhance the coarse segmentation information with less computation expense. Furthermore, they extended the network's gradient flow and used the low-resolution feature information. Jeevakala et al. [14] proposed a Mask R-CNN approach driven with U-Net to detect and segment the Internal Auditory Canal (IAC) and its nerves. In this method, the U-Net segmented the structure related information of IAC and its nerves by learning its features.

However, the variants' structure of U-Net suffers from redundant information. More and more network structures have been proposed and applied to image analysis [15].

Fisher and Koltun [16] proposed a new convolutional network module which used dilated convolutions. This module could aggregate multiscale contextual information systematically without losing resolution. Recently, dilated convolution is increasingly applied to medical images. Wolterink et al. [17] proposed a method to segment the myocardium and blood automatically in CMR of patient who has CHD by CNN. In the same year, Fisher et al. [18] developed a convolutional network module specifically for intensive prediction which used extended convolution to systematically aggregate multiscale context information without loss of resolution. Residual network (ResNet) was proposed in 2016, which added skip connections to each convolution layer for 2D image classification tasks. In addition, this architecture has been extended to 3D volumetric segmentation [19–21]. Huang et al. proposed the DenseNet with  $L(L+1)/2$  direct connections, which improved ResNet. It can strengthen feature propagation and reuse all features. After this improvement, Jégou [22] proposed a 2D fully convolutional DenseNet for semantic segmentation. In the same year, Yu et al. [23] proposed the DenseVoxNet; this network extended the deep residual learning in 2D image recognition tasks into 3D, which could simplify network training, reduce the parameters, and add auxiliary paths to enhance gradient propagation. However, there were no direct connections between the dense blocks and the final prediction layer. DenseVoxNet may not be able to appropriately capture multiscale contextual information useful for accurate segmentation. The correlation of adjacent images or frames should be effectively exploited for improving the accuracy of the target tasks which involves 3D volumetric data. Therefore, more and more methods have been proposed to use 3D features for biomedical volumetric data [24–29]; for example, Hosseini-Asl et al. [30] proposed a deep supervised adaptive 3D CNN, which could automatically extract and recognize the characteristics of Alzheimer's disease and capture the changes caused by Alzheimer's disease, such as the size of ventricle, the shape of the hippocampus, and the thickness of cortex. Dou et al. [31]

proposed a 3D fully convolutional network, called 3D Deeply Supervised Network (DSN), equipped with a deep supervision mechanism. This method has obtained good results in two tasks: liver segmentation of 3D CT scan, and whole heart and large blood vessels segmentation of 3D MRI. Previous CNN expresses dependencies between different image regions through convolution. Convolution operators have local receptive fields, so processing long-range dependencies goes through multiple convolutional layers, which may prevent learning about long-term dependencies. While it is possible to increase the representational capacity of the network by increasing the size of the convolutional kernels, the computational and statistical efficiency gained by using local convolutional structures are lost. However, self-attention [32–34] can exhibit a better balance between the ability to model long-range dependencies and computational efficiency. However, it is still a challenging task for CNNs to segment the important organs from 3D medical images due to the complexity of 3D structures, the difficulty of voxelized grid optimization, and the insufficiency in training samples.

Dou et al. [35] proposed Pnp-AdaNet using the method of adversarial learning, which could adapt to medical images of different modalities through plug-and-play modules. In another experiment, Dou et al. [36] constructed a domain adaptation module (DAM) to map the target region to features that were spatially aligned with the source domain region. The domain critic module (DCM) was responsible for distinguishing the feature spaces of the two domains. Then these two modules were optimized via an adversarial loss without using any target domain label. They trained the network using MRI, used it to segment CT images, and finally achieved certain results. The experiments done by Schlemper et al. [37] showed that using a grid-like attention mechanism in CT images might achieve better results. Shi et al. [38] proposed Bayesian VoxDRN for segmenting the entire heart from 3D MRI. Bayesian VoxDRN could predict voxel class labels by measuring the uncertainty of the model. During the test, it was realized by sampling based on Monte Carlo to generate a posteriori distribution of voxel labels. The attention mechanism was first applied to the text field. When the improved attention mechanism was applied to image processing, very good results were achieved. Liu et al. [39] proposed a novel medical image super-resolution method based on dense neural network and blended attention mechanism to address the problem that medical image would suffer from severe blurring caused by the lack of high-frequency details in the process of image super-resolution reconstruction. Kaul et al. [40] joined the attention tool to CNNs using feature maps generated by a separate convolutional autoencoder. This attention architecture was well suited for incorporation into deep convolutional networks. The results showed that this attention architecture was better than U-Net and residual variant.

### 3. Methods

**3.1. The Architecture of SparseVoxNet.** The architecture of SparseVoxNet proposed in this paper is shown in Figure 1. It improves U-Net which includes upsampling and

downsampling processes to implement end-to-end training. The padding is used for keeping the feature-map sizes constant in every sparse block, because the sparse block is not applicable when the feature maps have different sizes. Therefore, in each sparse block, the first 4 layers use ordinary convolutions, and the last 3 layers use dilated convolutions. The hole sizes are 2, 3, and 5. The spatial self-attention mechanism is added after the original feature map of data to strengthen the more important features in the original feature map. In the final deconvolution layer, instead of using a fully connected layer, three  $1 \times 1 \times 1$  convolution layers and softmax layer are used to obtain the segmented final label map. A dropout layer with a coefficient of 0.2 is added after each convolution layer to enhance the generalization ability of network.

Inspired by DenseNet, the black dotted line in SparseVoxNet in Figure 1 represents a skip connection. The image is segmented once by deconvolution on the skip connection. The network will converge faster and the accuracy rate will be higher due to the skip connection. The first segmented image will perform better on edge segmentation, because the shallow neural network loses less information through convolution and gets more edge information. The result is a fine grained segmentation. The result of the second segmented image is better in overall segmentation, which is coarse grained segmentation. Deep neural network features are high-level abstract features, which is really helpful when extracting the segmented central area of the entire tissue. The final segmentation result is determined by the voting of multiple segmentation results of different cropped input data on a single voxel point. The downsampling process of U-Net is replaced with sparse blocks, and the two deconvolutions are equivalent to the upsampling process.

Furthermore, we calculate the number of parameters for each layer in the SparseVoxNet shown in Table 1. Table 1 shows the parameters of 4 convolution layers, 2 deconvolution layers, 2 sparse blocks, a spatial attention mechanism layer, and a skip connection layer. Among them, the 4 convolution layers are represented by Conv\_n, the 2 deconvolution layers are represented by Deconv\_n, and the 2 sparse blocks are represented by Sparse Block\_n. We also show the convolution kernel and stride of each layer in Table 1. Note that each row in Table 1 corresponds to each layer in Figure 1.

**3.2. Sparse Block.** DenseNet has denser connections compared to ResNet, which makes the consumption of hardware resources very high. Therefore, we propose a sparse network structure to change the way of feature reuse while keeping feature reuse and skip connection characteristics unchanged. The sparse block which we propose reduces the number of connections, just having direct connections between any two layers with the same feature-map size, referred to as full skip connection, but the effect of sparse block is similar to dense block. The input of transition layer is as follows:

$$[T_0, T_1, T_2, T_3, T_4] = [T_0, H_1(T_0), H_2(T_1), H_3(T_2), H_4(T_3)]. \quad (1)$$



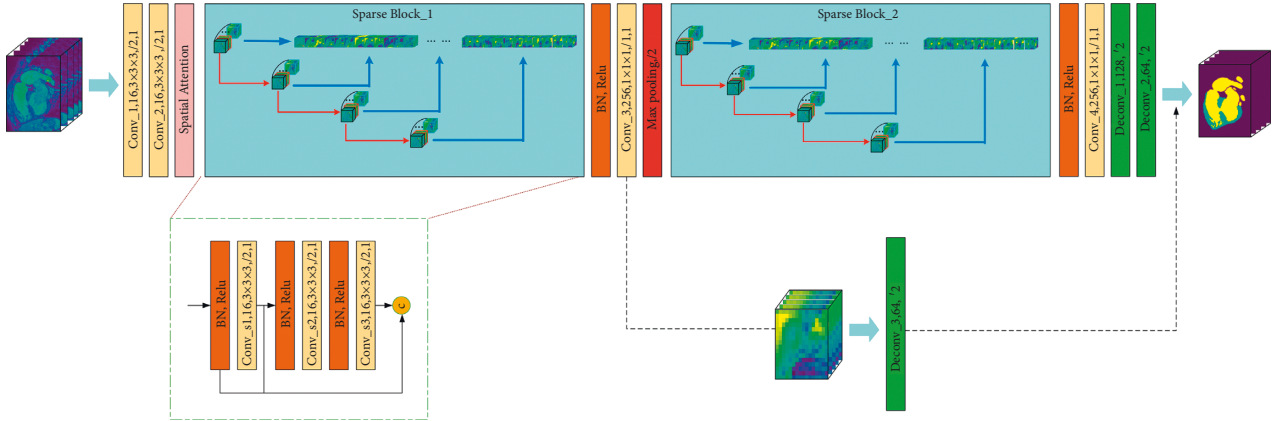


FIGURE 1: An overview of the proposed SparseVoxNet, with intermediate feature volumes. The light blue and dark blue areas of the slice represent the blood pool and myocardium. The dark blue and black areas belong to the background. The blue, yellow, and dark purple of segmented result represent the myocardium, blood pool, and background, respectively. There are two sparse blocks in this network. The black dotted line at the bottom right represents a skip connection.

TABLE 1: Our 3D convolutional model.

Input image		Output				Layer (type)		Stride	Kernel	Parameters	
64	64	64	1	32	32	32	16	Conv_1 (convolution)	2	3	448
32	32	32	16	16	16	16	16	Conv_2 (convolution)	2	3	6928
16	16	16	16	16	16	16	16	Spatial attention	2	1	816
16	16	16	16	16	16	16	100	Sparse Block_1 (sparse block)	1	3	43300
16	16	16	100	16	16	16	100	Conv_3 (convolution)	1	1	10100
16	16	16	100	16	16	16	184	Sparse Block_2 (sparse block)	1	3	496984
16	16	16	184	16	16	16	64	Conv_4 (convolution)	1	1	11840
16	16	16	64	32	32	32	64	Deconv_1 (deconvolution)	2	4	262208
32	32	32	64	64	64	64	64	Deconv_2 (deconvolution)	2	4	262208
16	16	16	100	64	64	64	64	Skip connection	1	1	6464

where the input of  $H_1$  is  $T_0$ , the input of  $H_2$  is  $T_0 + T_1$ , and so on.

The feature maps of different receptive fields are referred to as different scales. It is found that the nonlinear combination of the features of different scales is not better than the linear combination. Inspired by the U-Net network structure, composite expression features are constructed by directly stacking feature maps of different scales.

DeletedUnlike U-Net, the improved network structure uses deconvolution to replace the upsampling process, which reduces the loss of information during the conversion process. In DenseNet, the network connections of the previous layer and the latter layer are too dense, which can easily cause overfitting. The sparse network can solve this problem. The network's feature expression ability is greatly enhanced, and there is no vanishing gradient.

**3.3. 3D Dilated Convolution.** Dilated convolution has one more hyperparameter than traditional convolution, called dilation rate. Dilated convolution adds holes to the standard convolution kernel. In this paper, we extend the dilated convolution to 3D data, and mix the traditional convolution and dilated convolution. Referring to the DenseVoxNet, we use 4 layers of  $3 \times 3 \times 3$  traditional convolution and 3 layers

of dilated convolution with 2, 3, and 5 holes. The 4 layers of traditional convolution can extract the local features of the image, and the 3 layers of dilated convolution expand the reception field of the feature exponentially to capture the potential relationship between long distance features. We only use 7-layer convolution to make the reception field reach  $26 \times 26 \times 26$ .

**3.4. Spatial Self-Attention Mechanism.** In both the computer vision tasks and the natural language processing tasks, the dependencies between long distance features are difficult to capture. In serialization tasks, recurrent and recursive neural networks are major means to capture long-range dependencies. In convolutional neural networks, large reception fields are formed by superposing multiple convolution operations. Currently, there are no specific methods to capture long-range features. Convolution and cyclic operators have the following disadvantages: (1) being too inefficient, (2) easily producing gradient disappearance, (3) difficulty of passing information back and forth between long ranges.

Inspired by the nonlocal mean filtering for images, Wang et al. [41] proposed nonlocal block for capturing long-range dependencies, which is a self-attention mechanism.

Nonlocal block ignores the Euclidean distance and calculates the relationship between two positions directly. Actually, it calculates the generalized autocorrelation matrix of features. However, the calculation efficiency is relatively high. Because after adding nonlocal operators, it is not necessary to stack too deep convolution operations for achieving the network's fitting ability. Furthermore, it does not change the size of input data and can be easily embedded in the network, so the spatial self-attention model is added in front of the first sparse block. We apply the self-attention mechanism, proposed by Zhang et al. [42], in this paper. The nonlocal block is embedded in the 3D network, which is defined as follows:

$$y_i = \frac{1}{C(x)} \sum_{\forall j} f(x_i, x_j)g(x_j), \quad (2)$$

where  $i$  is a 3D coordinate meaning the position index of input data,  $j$  is the index of all possible positions,  $x$  is the input data,  $f$  is an autocorrelation calculation function, which can calculate the correlation between  $i$ -th position and  $j$ -th position, and  $g$  is a unary mapping function. The  $1 \times 1 \times 1$  convolution is used for ascending dimension and fusing the multichannel feature in the experiments, and finally  $C(x)$  is used for normalization. Using multiple  $1 \times 1 \times 1$  convolution kernels in the attention model can not only achieve cross-channel interaction and information integration, but also reduce or increase the number of channels. People begin to pay attention to the  $1 \times 1 \times 1$  convolution because of the network structure proposed by Lin [43]; this convolution connects two full connection layers for fusing the features linearly. After that, in Google's Inception-v4 [44] network structure,  $1 \times 1 \times 1$  convolution is used in the inception module for dimensionality reduction or ascending dimension. Inspired by this advantage, in this paper, the  $1 \times 1 \times 1$  convolution kernel is used to reduce the original input data dimensionality, calculate the spatial autocorrelation relationship, and then ascend the dimension of data. The different weights calculated are added back to the original data and then regularized to describe the influence on features of voxel points in different spatial positions.

## 4. Experiments and Results

**4.1. Dataset.** Radiobiological images mainly have six data formats. The NIFTI (Neuroimaging Informatics Technology Initiative) is one of them. The data format used in this paper is NIFTI. This format contains two affine coordinates, so that it can associate the physical index of voxels with its actual spatial location. The HVSMR 2016 dataset is used to evaluate the algorithm and network structure. HVSMR 2016 has a total of 10 cardiac magnetic resonance 3D scans for training and 10 scans for testing. All training sets of cardiac MRIs are from patients with CHD, including annotations of myocardium and large blood vessels.

Due to the large difference in intensity between different images, the cardiac MR images are all normalized. After normalization, the mean and unit variance are 0. To leverage

the limited training data, simple data augmentation was employed to enlarge the training data. The augmentation operations include the rotation and cropping. The original training set is divided into three parts, namely, the training set, the validation set, and the testing set. The cross-validation method is used for parameter training. We use 70% of the images for training and 30% for testing. Then, we compare and briefly discuss the experimental results.

**4.2. Evaluation Metrics.** Medical image segmentation is an important step of medical image processing. However, it is difficult to select accurate evaluation index to evaluate the quality of segmentation by comparing segmented medical images. The following three metrics are used in this paper for measuring the results of segmentation.

**4.2.1. Dice Coefficient.** Dice coefficient is widely used for verifying the effect of 3D medical image segmentation. The core idea is to ensure a high recall and precision. Compared with the evaluation method of directly computing the difference between the automatic segmentation results and the original data labels, using Dice coefficient can better characterize the segmentation effect. Dice coefficient is defined as follows:

$$\text{Dice} = \frac{2|G \cap R|}{|R| + |G|} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}}, \quad (3)$$

where  $G$  is the segmentation result of ground truth, which is the labeled testing data.  $R$  is the automatic segmentation result of testing data. TP, FP, and FN represent true positives, false positives, and false negatives, respectively, for each class. Ideally, the template of segmentation result and the template of label data completely overlap, which means  $R = G$ , and the absolute value of the Dice coefficient is 1.

**4.2.2. Average Symmetric Surface Distance.** Average symmetric surface distance (ADB) is defined as follows: for a single voxel point, if one or more voxel points within its 18-neighborhood are not elements of the object, they are regarded as surface voxel points. For each surface voxel point in the  $R$ , we calculate the Euclidean distance between it and the nearest surface voxel point of the real label  $G$ . Similarly, perform the same calculation for each surface voxel point in the  $G$ .  $S(R)$  represents the surface voxel point set of  $R$ . The distance from any voxel point  $v$  to  $S(R)$  is defined as  $d(v, S(R)) = \min_{s_R \in S(R)} \|v - s_R\|$ . Based on this formula, the average symmetrical surface distance is defined as follows:

$$\text{ADB} = \frac{1}{|S(G)| + |S(R)|} \times \left[ \sum_{s_R \in S(R)} d(s_R, S(G)) + \sum_{s_G \in S(G)} d(s_G, S(R)) \right]. \quad (4)$$



Many segmentation boundary evaluation metrics are constructed based on this distance formula, which measures the boundary difference between the segmentation result and the ground truth by calculating the voxel surface distance. The larger the value of ADB, the more dissimilar the segmentation boundary is. When the boundary of the segmentation result matches the ground truth exactly, the value of ADB is 0.

**4.2.3. Hausdorff Distance.** Based on the ADB, when using the maximum symmetric distance, the metric is known as Hausdorff distance, which is defined as follows:

$$\text{hausdorff}(R, G) = \max_{r \in R} \left\{ \min_{g \in G} \{d(r, g)\} \right\}, \quad (5)$$

where  $d(r, g)$  represents the distance between points  $r$  and  $g$ ; that is, the set consists of the shortest distance (usually expressed in Euclidean distance) from all points in the predicted segmentation set  $R$  to any point in the real label set  $G$ , and the maximum distance is selected from this set as the Hausdorff distance between the two sets  $R$  and  $G$ . This distance and the symmetrical surface distance both describe the similarity of the contour. The larger the absolute value, the less similar the segmentation.

**4.3. Training.** In the experiments, all weights are randomly initialized by the Gaussian distribution with  $\mu = 0$ ,  $\theta = 0.01$ , and the stochastic gradient descent optimization algorithm is used. Batch size is set to 8. In order to reduce the model overfitting and speed up the convergence rate, the weight attenuation is 0.0005, and the momentum is set to 0.9, which is often used to speed up training, while making it easier to jump out of extreme points and avoid getting stuck in local optimal solutions. The drop rate is 0.2, and the initial learning rate is set to 0.01. If the learning rate is too low, the training period is too long, and the high learning rate will cause the model to be unstable and never converge. Our algorithms were trained and tested on the Dual RTX 2080 Ti GPU.

The polynomial decaying learning rate is used for ensuring the rapid convergence of the model during the initial training period and the stability of the model parameters in the later period. The initial learning rate is set relatively large, and the learning rate is reinitialized and decayed every 5000 steps. The attenuation coefficient of the learning rate is  $\delta = (1 - \text{iter}/\text{max\_iter})^{\text{power}}$ . After testing, the model stabilizes after 8000 iterations. The input data of SparseVoxNet consists of 8 groups of  $64 \times 64 \times 64$  heart MRIs, which are cropped randomly in the same axis direction.

Multiple sets of comparative experiments and ablation experiments are designed to verify the effect of the improved method on segmentation. In the experiments, we compared our method with the traditional methods and other deep learning methods, and also compared the network only with the mixed dilated convolution and the network only with the attention mechanism and DenseVoxNet.

**4.3.1. Ablation Study.** We conduct ablation experiments to verify the importance of 3D dilated convolution and spatial self-attention mechanism in exploiting multiscale features. The results are presented in Table 2.

When we just add the mixed 3D dilated convolution to the model, we define this model as SparseVoxNet-D. The Dice coefficient of myocardium and blood pool gets the best results, 82.4% and 91.6%, respectively. It verifies our conjecture: dilated convolutions can exponentially expand receptive fields to obtain multiscale information without losing resolution or coverage, especially for structures with a small size or irregular boundary, such as the cardiac myocardium structures. Since the receptive field expansion speed of the dilated convolution depends on the number of holes in the dilated convolution, although the more holes will contribute a larger receptive field, the pixels in the large receptive field are not necessarily related to the current convolution. In other words, the larger receptive field is not the better. Local perception can better capture local features. Global perception can better capture the relationship characteristics of pixels at different locations. Hence, we mix the 4 layers of traditional convolution and 3 layers of dilated convolution and define the different dilation rates of dilated convolutions to better capture features.

When we just add the spatial self-attention mechanism to the model, we define this model as SparseVoxNet-S. It can be seen that ADB and Hausdorff distance of blood pool and myocardium achieve better performance than DenseVoxNet, the Hausdorff distance of myocardium outperforms DenseVoxNet by around 3.0%, and the Hausdorff distance of blood pool outperforms DenseVoxNet by around 4.8%. This indicates that with the spatial self-attention mechanism, the segmented images have been brought closer to the target domain successfully, because the self-attention in our model is complementary to the convolution for capturing long-range, global-level dependencies occurring in cardiac structure. The advantages of the attention mechanism are as follows: (a) few parameters; (b) fast calculation; (c) capturing long-range features. The problem applied in this paper is a small sample training process, so when the spatial self-attention mechanism is removed, the segmentation result is not ideal, which means the long-range features cannot be extracted efficiently. We use both dilated convolution and spatial self-attention mechanism to capture long-range features, because the method based on dilated convolution obtains information from a small number of surrounding points and cannot form dense context information. The spatial self-attention mechanism makes a single feature in any location perceive the features of all other locations, and can produce more powerful pixel-level representation capabilities. These observations demonstrate that the 3D dilated convolution and the spatial self-attention indeed play a meaningful role in exploiting multiscale features.

**4.4. Results.** There are segmentation results on three training images shown in Figure 2. These three slices come from different patients. The data whose indexes are 60 in the

TABLE 2: Results of ablation study. Bold results are the best ones.

Method	Myocardium			Blood pool		
	Dice (%)	ADB	Haus.	Dice (%)	ADB	Haus.
SparseVoxNet-D	<b>82.4</b>	0.922	5.385	<b>91.6</b>	1.073	7.736
SparseVoxNet-S	80.7	<b>0.853</b>	<b>5.075</b>	91.4	<b>0.951</b>	<b>5.004</b>
DenseVoxNet	79.2	0.943	7.175	89.48	0.955	9.608

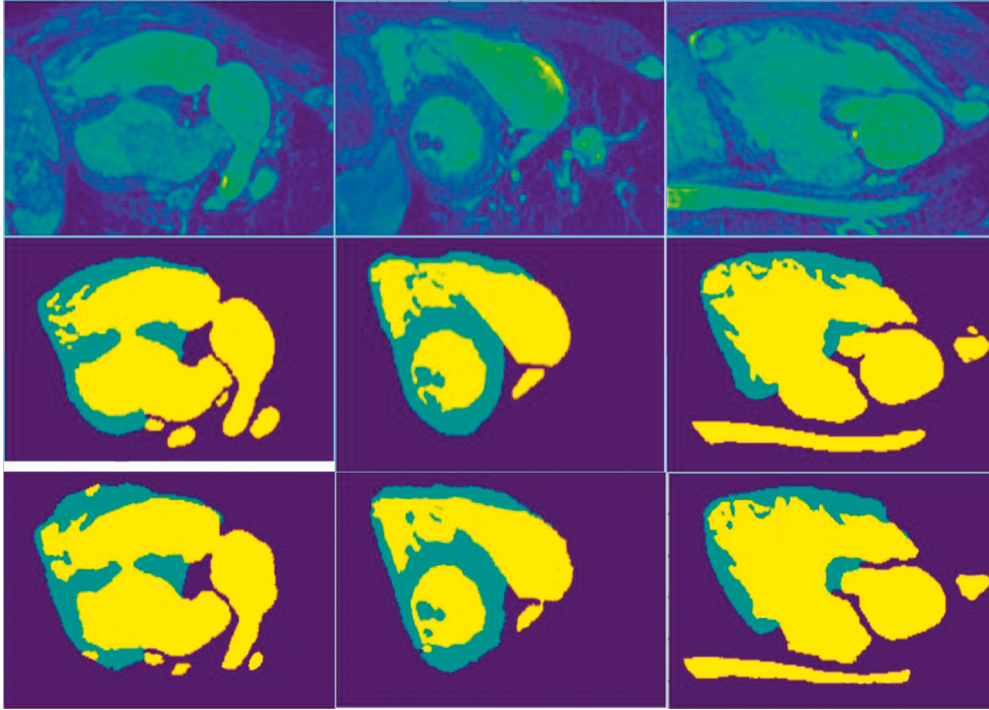


FIGURE 2: Segmentation results on three training images.

sample dataset have the same coronal plane view in the same dimension. The light blue and dark blue areas of the image in the first line represent the blood pool and myocardium; the dark blue and black areas belong to the background. The images in the second line are labeled, corresponding to the myocardium and blood pool in the first line of images. The third line is the results of automatic segmentation by the method proposed in this paper, where blue, yellow, and dark purple represent the myocardium, blood pool, and background, respectively. It can be seen from Figure 2 that although the cardiac structure of different patients in the training set is quite different, the method we proposed can still successfully calibrate the myocardium and blood pool from low contrast cardiac MRI, which proves that this method has a good enough fitting ability to the original data. However, there are still some disadvantages. In the first auto-segmentation result, the myocardium in the lower left corner is partially divided. In the second result, the background appeared in the myocardium. In the third result, there is extra myocardium in the upper right corner, which shows that deep learning has the ability to perceive most data features, but it does not have reasonable logical reasoning

capabilities. Human segmentation will not produce these subtle logical errors.

Figure 3 shows segmentation results on three testing images. The data extraction method is the same as above. By observing the results, we can see that the method proposed in this paper also has good generalization effect on unlabeled data. However, when using the gradient descent algorithm, it is easy to fall into the local optimum and cause overfitting, because of the huge number of parameters.

*4.5. Discussion.* The comparison of the results between the method we proposed and other six methods is shown in Figure 4. They are mainly ranked according to the Dice coefficient. The figure also shows the auxiliary reference indexes, such as ADB and symmetric Hausdorff distance. The first three are traditional methods, such as manually extracting features and using hidden Markov random fields, and the other deep learning methods are on the HVSMR 2016 Challenge dataset. According to Figure 4, the Dice coefficient of blood pool in all methods is higher than that of myocardium, suggesting that the segmentation of blood pool is

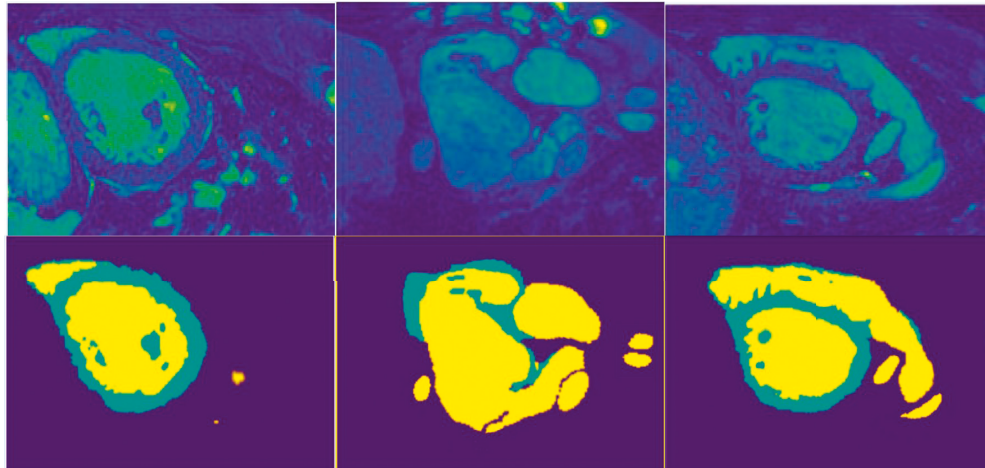


FIGURE 3: Segmentation results on three testing images: The three slices in the first line come from different patients. Images in the second line are the results of automatic segmentation by the method proposed in this paper.

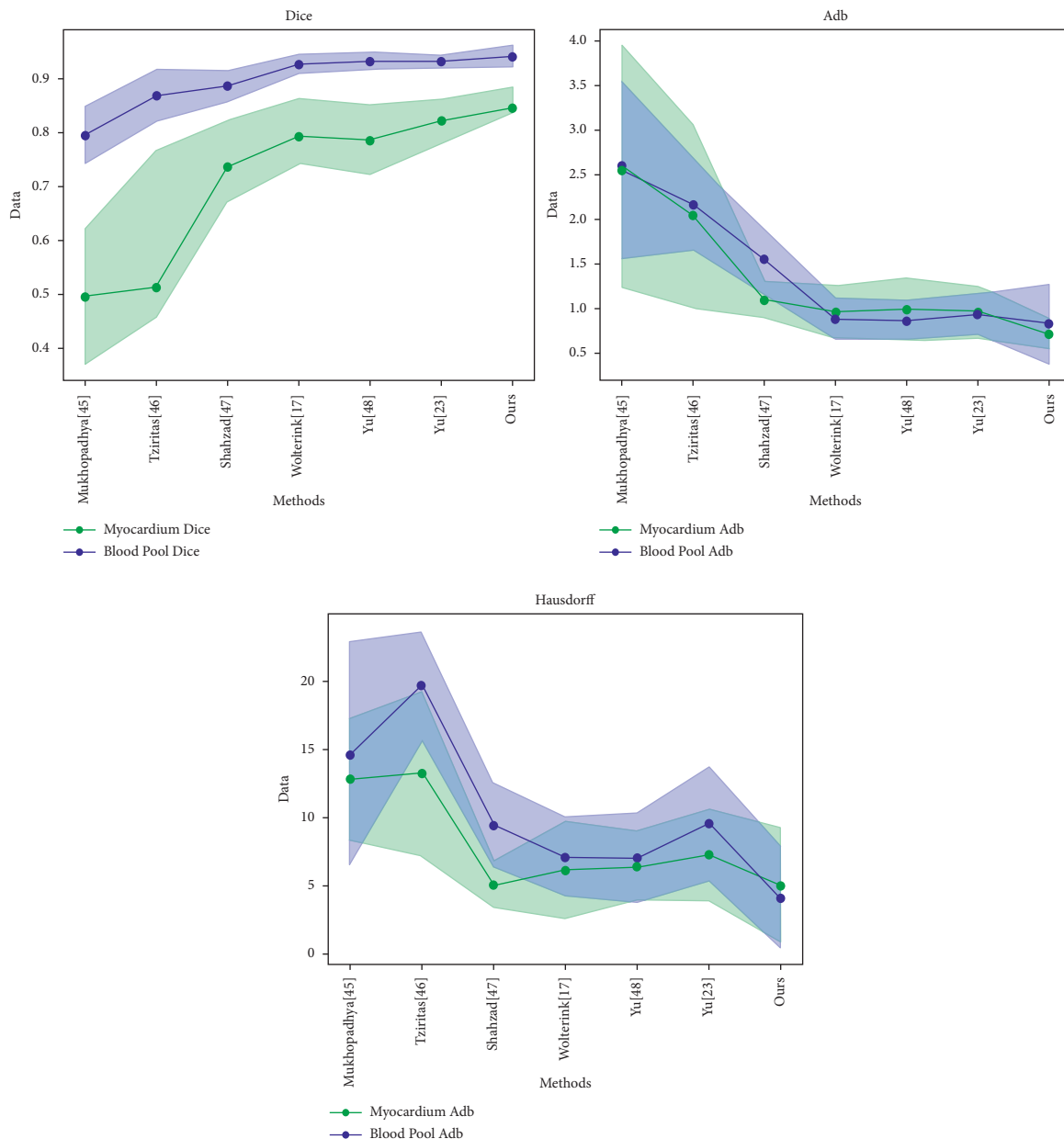


FIGURE 4: Comparison of experimental results between the improved method and other methods.

TABLE 3: Comparison of experimental results between the improved method and the 3D methods. Bold results are the best ones.

Method	Myocardium			Blood pool		
	Dice (%)	ADB	Haus.	Dice (%)	ADB	Haus.
3D U-Net [11]	69.4	2.596	12.796	79.4	2.550	14.634
V-Net [26]	70.3	2.367	10.624	81.9	2.435	12.539
VoxResNet [19]	77.4	2.041	13.199	86.7	2.157	19.723
DenseVoxNet [23]	79.2	0.943	7.175	89.48	0.955	9.608
Ours	<b>84.5</b>	<b>0.721</b>	<b>5.027</b>	<b>94.0</b>	<b>0.831</b>	<b>4.102</b>

relatively easier due to the ambiguous borders of the myocardium in the low-resolution MRIs. Regarding the segmentation of myocardium, the method we proposed achieves the best performance with the Dice; i.e., the ranking metric in the challenge,  $0.861 \pm 0.024$ , outperforms the second one by around 4%. The best result also has been achieved in blood pool segmentation with Dice; the ranking metric in the challenge,  $0.94 \pm 0.016$ , demonstrates that our sparse connected network has the capability to tackle hard cardiovascular segmentation problem. The ADB and Hausdorff distance of our method also achieved the best performance.

The results of other 3D MRI segmentation methods are mainly shown in Table 3. Firstly, the experimental parameters are compared, and the method proposed in this paper needs the least parameters. The sparse block and dilated convolution can achieve a good fitting effect with the participation of such a small number of parameters, thanks to the introduction of the attention model. The feature expression ability of sparse block will not be better than dense block in many cases, but the problem applied in this paper is medical segmentation and a small sample training process, so the sparse block can fit and generalize the data well with a small number of parameters, and the exponentially increasing receptive field provided by the dilated convolution reduces the convolution operations. The attention mechanism can well capture the features to strengthen the generalization ability of the network. Because of the small number of parameters, the amount of calculation is reduced, and the model's convergence rate is also fast.

Comparing the cross-entropy loss of DenseVoxNet and SparseVoxNet with sparse block and dilated convolution, we can find that the network only using the mixed dilated convolution can converge faster and reach lower loss values, which proves that the improved sparsely connected network structure can reduce the calculation amount and improve the efficiency and that the method of extracting long-range features by hybrid 3D dilated convolution is suitable for medical images. It has better ability to represent features and fit data.

The comparison shows that the time of one iteration of DenseVoxNet (forward and backward propagation of the network) is 0.113 s, the time of SparseVoxNet-D is 0.045 s, and the time of SparseVoxNet is 0.049 s. The proposed method has a great improvement in efficiency.

## 5. Conclusion

In this paper, we propose a novel and efficient 3D sparse convolutional network to segment blood pool and myocardium from 3D cardiac magnetic resonance images. This

method can eliminate redundant calculations and reduce model parameters and the risk of overfitting training data on small samples. The spatial self-attention mechanism can optimize the expression ability of feature maps, and the sparse blocks can reduce the convolutional network depth. The work in this paper is an accurate pixel-level classification. Moreover, we achieve competitive results in comparison with existing methods. The proposed method can provide comprehensive information for doctors to make diagnoses of CHD.

## Data Availability

The data used in this study could be accessed upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported by National Natural Science Foundation of China (No. 61872162 and 82071995) and Natural Science Foundation of Jilin Province under Grant No. 20200201292JC.

## References

- [1] M. Bakator and D. Radosav, "Deep learning and medical diagnosis: a review of literature," *Multimodal Technologies and Interaction*, vol. 2, no. 2, p. 47, 2018.
- [2] A. S. Lundervold and A. Lundervold, "An overview of deep learning in medical imaging focusing on MRI," *Zeitschrift für Medizinische Physik*, vol. 29, no. 2, 2018.
- [3] X. Zhuang, "Challenges and methodologies of fully automatic whole heart segmentation: a review," *Journal of Healthcare Engineering*, vol. 4, no. 3, pp. 371–407, 2015.
- [4] W. G. Hatcher and W. Yu, *A Survey of Deep Learning: Platforms, Applications and Emerging Research Trends*, p. 1, IEEE Access, USA, 2018.
- [5] O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, Germany, 2015.
- [6] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 4, pp. 640–651, 2014.
- [7] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern*

- Recognition 14 (CVPR)*, IEEE Computer Society, Hawaii, USA, 2017.
- [8] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Nevada, 2016.
  - [9] T. X. Pham, P. Siarry, and H. Oulhadj, "A multi-objective optimization approach for brain MRI segmentation using fuzzy entropy clustering and region-based active contour methods," *Magnetic Resonance Imaging*, vol. 61, pp. 41–65, 2019.
  - [10] D. U. Hongwei, K. Shao, F. Bao et al., "Automated coronary artery tree segmentation in coronary CTA using a multi-objective clustering and toroidal model-guided tracking method," *Computer Methods and Programs in Biomedicine*, vol. 9963, Article ID 105908, 2020.
  - [11] Ö. Cicek, A. Abdulkadir, S. Lienkamp, T. Brox, and O. Ronneberger, "3D U-net: learning dense volumetric segmentation from sparse annotation," in *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2016. MICCAI 2016. Lecture Notes in Computer Science*, S. Ourselin, L. Joskowicz, M. Sabuncu, G. Unal, and W. Wells, Eds., vol. 9901, Berlin, Germany, Springer, 2016.
  - [12] M. Habijan, H. Leventić, I. Galić, and D. Babin, "Whole heart segmentation from CT images using 3D U-net architecture," in *Proceedings of the International Conference System, Signals Image Process (IWSSIP)*, pp. 121–126, Osijek, Croatia, June 2019.
  - [13] X. Ding, Y. Peng, C. Shen, and T. Zeng, "Cab U-net: an end-to-end category attention boosting algorithm for segmentation," *Computerized Medical Imaging and Graphics*, vol. 84, Article ID 101764, 2020.
  - [14] S. Jeevakala, C. Sreelakshmi, K. Ram, R. Rangasami, and M. Sivaprakasam, "Artificial intelligence in detection and segmentation of internal auditory canal and its nerves using deep learning techniques," *International Journal of Computer Assisted Radiology and Surgery*, vol. 15, no. 11, pp. 1859–1867, 2020.
  - [15] G. Litjens, T. Kooi, B. E. Bejnordi et al., "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, 2017.
  - [16] Y. Fisher and V. Koltun, *Multi-Scale Context Aggregation by Dilated Convolutions*, 2016.
  - [17] J. M. Wolterink, T. Leiner, M. A. Viergever, and I. Išgum, "Dilated convolutional neural networks for cardiovascular MR segmentation in congenital heart disease," in *Proceedings of the International Workshop on Reconstruction and Analysis of Moving Body Organs International Workshop on Whole-Heart and Great Vessel Segmentation from 3D Cardiovascular MRI in Congenital Heart Disease*, pp. 95–102, Athens, Greece, 2017.
  - [18] Y. Fisher, V. Koltun, and T. Funkhouser, "Dilated residual network," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 472–480, Honolulu, HI, USA, 2017.
  - [19] H. Chen, Q. Dou, L. Yu, Q. Jing, and H. Pheng-Ann, "VoxResNet: deep voxelwise residual networks for brain segmentation from 3D MR images," *NeuroImage*, vol. 170, pp. 446–455, 2017.
  - [20] L. Yu, X. Yang, H. Chen, J. Qin, and P. A. Heng, "Volumetric ConvNets with mixed residual connections for automated prostate segmentation from 3D MR images," in *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, pp. 66–72, AAAI Press, 2017.
  - [21] A. Fakhry, T. Zeng, and S. Ji, "Residual deconvolutional networks for brain electron microscopy image segmentation," *IEEE Transactions on Medical Imaging*, vol. 36, no. 2, pp. 447–456, 2017.
  - [22] S. Jégou, M. Drozdal, D. Vazquez, A. Romero, and Y. Bengio, "The one hundred layers tiramisu: fully convolutional denets for semantic segmentation," in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Honolulu, HI, USA, 2017.
  - [23] L. Yu, J.-Z. Cheng, Q. Dou et al., "Automatic 3D cardiovascular MR segmentation with densely-connected volumetric convNets," in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 287–295, Canada, 2017.
  - [24] M. F. Stollenga, W. Byeon, M. Liwicki, and J. Schmidhuber, "Parallel multi-dimensional LSTM, with application to fast biomedical volumetric image segmentation," *Advances in Neural Information Processing Systems*, vol. 2, pp. 2998–3006, 2015.
  - [25] Q. Dou, H. Chen, L. Yu et al., "Automatic detection of cerebral microbleeds from MR images via 3D convolutional neural networks," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1182–1195, 2016.
  - [26] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-Net: fully convolutional neural networks for volumetric medical image segmentation," in *Proceedings of the 2016 Fourth International Conference on 3D Vision (3DV)*, pp. 565–571, IEEE, Stanford, California, 2016.
  - [27] S. Andermatt, S. Pezold, and P. Cattin, "Multi-dimensional gated recurrent units for the segmentation of biomedical 3D-data," in *Proceedings of the International Workshop on Large-Scale Annotation of Biomedical Data and Expert Label Synthesis International Workshop on Deep Learning in Medical Image Analysis*, pp. 142–151, Athens, Greece, 2016.
  - [28] K. Kamnitsas, C. Ledig, V. F. J. Newcombe et al., "Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation," *Medical Image Analysis*, vol. 36, pp. 61–78, 2017.
  - [29] J. Dolz, C. Desrosiers, and I. B. Ayed, "3D fully convolutional networks for subcortical segmentation in MRI: a large-scale study," *NeuroImage*, vol. 170, pp. 456–470, 2017.
  - [30] E. Hosseini-Asl, R. Keynto, and A. El-Baz, "Alzheimer's disease diagnostics by a deeply supervised adaptable 3D convolutional network," in *Proceedings of the 2016 IEEE International Conference on Image Processing (ICIP)*, Phoenix, AZ, USA, 2016.
  - [31] Q. Dou, H. Chen, Y. Jin, L. Yu, J. Qin, and P.-A. Heng, "3D deeply supervised network for automatic liver segmentation from CT volumes," in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, Athens, Greece, 2016.
  - [32] J. Cheng, L. Dong, and M. Lapata, *Long Short-Term Memory Networks for Machine Reading*, EMNLP, 2016.
  - [33] A. P. Parikh, O. Tackström, D. Das, and J. Uszkoreit, *A Decomposable Attention Model for Natural Language Inference*, EMNLP, 2016.
  - [34] A. Vaswani, N. Shazeer, N. Parmar et al., "Attention is all you need," 2017, <https://arxiv.org/abs/1706.03762>.
  - [35] Q. Dou, C. Quyang, C. Chen et al., "PnP-AdaNet: plug-and-play adversarial domain adaptation network with a benchmark at cross-modality cardiac segmentation," 2018, <https://arxiv.org/abs/1812.07907>.
  - [36] Q. Dou, C. Quyang, C. Chen, H. Chen, and P.-A. Heng, "Unsupervised cross-modality domain adaptation of convnets

- for biomedical image segmentations with adversarial loss,” 2018, <https://arxiv.org/abs/1804.10916>.
- [37] J. Schlemper, O. Oktay, M. Schaap et al., “Attention gated networks: learning to leverage salient regions in medical images,” *Medical Image Analysis*, vol. 53, pp. 197–207, 2019.
- [38] Z. Shi, G. Zeng, L. Zhang et al., “Bayesian VoxDRN: a probabilistic deep voxelwise dilated residual network for whole heart segmentation from 3D MR images,” in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 569–577, Spain, 2018.
- [39] K. Liu, Y. Ma, and H. Xiong, *Medical Image Super-Resolution Method Based on Dense Blended Attention Network*, 2019.
- [40] C. Kaul, S. Manandhar, and N. Pears, “FocusNet: an attention-based fully convolutional network for medical image segmentation,” in *Proceedings of the 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI)*, Italy, 2019.
- [41] X. Wang, R. Girshick, A. Gupta, and H. Kaiming, “Non-local Neural Networks,” in *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 7794–7803, Salt Lake City, UT, USA, 2018.
- [42] H. Zhang, I. Goodfellow, and D. Metaxas, “Self-attention generative adversarial networks” augustus odena proceedings of the 36th international conference on machine learning,” *PMLR*, vol. 97, pp. 7354–7363, 2019.
- [43] M. Lin, Q. Chen, and S. Yan, “Network In Network,” 2013, <https://arxiv.org/abs/1312.4400>.
- [44] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, “Inception-v4, inception-resnet and the impact of residual connections on learning,” in *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, California, USA, 2017.
- [45] A. Mukhopadhyay, “Total variation random forest: fully automatic MRI segmentation in congenital heart disease,” *Reconstruction, Segmentation, and Analysis of Medical Images*, vol. 10129, 2016.
- [46] G. Tziritas, “Fully-automatic segmentation of Cardiac images using 3-D MRF model optimization and substructures tracking,” in *Proceedings of the International Workshop on Reconstruction & Analysis of Moving Body Organs*, Springer International Publishing, Athens, Greece, 2016.
- [47] R. Shahzad, S. Gao, and Q. Tao, “Automated cardiovascular segmentation in patients with congenital heart disease from 3D CMR scans: combining multi-atlases and level-sets,” in *Proceedings of the International Workshop on Reconstruction & Analysis of Moving Body Organs*, Springer International Publishing, Athens, Greece, 2016.
- [48] L. Yu, X. Yang, J. Qin, and P. Heng, “3D FractalNet: dense volumetric segmentation for cardiovascular MRI volumes,” *Reconstruction, Segmentation, and Analysis of Medical Images*, vol. 10129, pp. 103–110, 2017.

## *Retraction*

# **Retracted: Picture Fuzzy Einstein Hybrid-Weighted Aggregation Operator and Its Application to Multicriteria Group Decision Making**

### **Computational Intelligence and Neuroscience**

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### **References**

- [1] G. Cao, "Picture Fuzzy Einstein Hybrid-Weighted Aggregation Operator and Its Application to Multicriteria Group Decision Making," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 6925670, 20 pages, 2022.



## Research Article

# Picture Fuzzy Einstein Hybrid-Weighted Aggregation Operator and Its Application to Multicriteria Group Decision Making

Guo Cao 

*School of Economics and Management, Changzhou Institute of Technology, Changzhou 213032, China*

Correspondence should be addressed to Guo Cao; caog@czu.cn

Received 16 December 2021; Accepted 31 March 2022; Published 14 May 2022

Academic Editor: Daqing Gong

Copyright © 2022 Guo Cao. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

As an extension of intuitionistic fuzzy sets (IFSs), picture fuzzy sets (PFSs) can better model and represent the hesitancy and uncertainty of decision makers' preference information. In this study, we propose a multicriteria group decision making (MCGMD) method based on picture fuzzy sets. We first define some basic picture Einstein operations with closed properties among PFSs on the basis of the Einstein  $t$ -norms and  $t$ -conorms. Then, utilizing the hybrid-weighted operator and the developed picture Einstein operations laws, we put forward a picture fuzzy Einstein hybrid-weighted aggregation operator for aggregating PFSs and discuss its several important properties. Furthermore, we present a new MCGMD method based on the proposed picture fuzzy Einstein hybrid-weighted aggregation operator. Finally, an example is conducted to validate the effectiveness of the proposed MCGMD method.

## 1. Introduction

Multicriteria group decision making (MCGDM) is one of the most important human activities [1–6]. Due to the complexity and vagueness of information in group decision making, however, it is usually difficult for decision makers to evaluate the alternatives by crisp numerical values [7, 8]. In some occasions, it can be more reasonable to give uncertain or fuzzy evaluation information represented by fuzzy numbers [9, 10], intuitionistic fuzzy sets (IFS) [11–15], neutrosophic set (NS) [16–18], Pythagorean fuzzy sets [19, 20], Fermatean fuzzy sets [21, 22], hesitant fuzzy sets [23–25], and so on. Based on the concept of fuzzy sets (FSs), Atanassov [26] first introduced the concept of IFS, which is an extension of Zadeh's FSs. Different from FSs, it does not require that the sum of the degrees of membership and non-membership of an element to be equal to one. IFSs have been successfully applied to multicriteria group decision making. However, there are some scenarios that cannot be represented by IFSs in some real-life group decision-making problems, such as the option of hesitation or remaining neutral. To overcome this drawback, Smarandache [27] proposed the concept of the NS, which has the degrees of

truth, falsity, and indeterminacy, respectively. Considering that it is difficult to utilize NSs to solve real-life scientific and engineering problems, Zhang and Sunderraman [28] proposed the concept of the single-valued neutrosophic set (SVNS). In this paper, Boran and Akay [11] also developed some set-theoretic operators and discussed their various properties. Compared with NS, it is assumed in a SVNS that each type of membership degree can take its values in the interval  $[0, 1]$ , and the sum of its truth-membership degree, indeterminacy-membership degree, and false-membership degree is less than or equal to 3. This hypothesis implies that its three types of membership degrees do not satisfy probabilistic independence. This non-restriction leads to dialetheist and paraconsistent information in real-life decision-making problems represented as SVNSs. In order to overcome these drawbacks of IFSs and NSs, Cuong and Kreinovich [29] proposed the concept of the picture fuzzy set (PFS) to address a fact such that human opinions involve several types of answers such as yes, abstain, no, or refusal and so on. The PFS is characterized by a positive membership function, a negative membership function, and a neutral membership function, and the sum of its three membership degrees is less than or equal to 1. Different from



NS, there is a restriction on the sum of the three types of membership degrees in PFS, which implies that they are dependent on each other.

The PFS is an extension of FS and IFS [30]. The core of PFS is its picture fuzzy value (PFV), which is composed of the degree of positive, negative, and neutral memberships. Similar to an intuitionistic fuzzy value (IFV) and NS, PFV is also a very effective tool to express inherent uncertainty or imprecision of decision makers' preference information in human decision-making processes. Recently, PFS has been successfully applied to a lot of areas. For example, Zhang et al. [23] introduced the concept of generalized picture distance measure and applied it to pattern recognition. Based on the correlation measure of Atanassov's IFSs, Singh [31] also proposed the concept of correlation for PFSs to solve bidirectional approximate reasoning systems problems. By combining picture composite cardinality with PSO, Thong and Son [32] proposed a novel automatic picture fuzzy clustering method for pattern recognition and knowledge discovery.

In the course of MCGMD with IFS, hesitant fuzzy set, and NS, the aggregation operators play a very important role [33]. Many scholars have introduced various aggregation operators for IFS, hesitant fuzzy set, and NS, such as intuitionistic fuzzy aggregation operators [34], hesitant fuzzy aggregation operators [35], interval-valued intuitionistic fuzzy aggregation operators [36], the neutrosophic fuzzy aggregation operators [37, 38] picture fuzzy aggregation operators [39–41], and so on. Note that most of the abovementioned aggregation operators were developed based on the triangular t-norm and t-conorm for the following reasons. For IFSs, they only involve membership degree and non-membership degree such that their sum is less than or equal to 1. Thus, triangular t-norms and t-conorms are appropriate in developing the intuitionistic fuzzy aggregation operators. As for SVNS, the sum of its truth-membership degree, indeterminacy-membership degree, and false-membership degree is less than or equal to 3. Hence, triangular t-norm or t-conorm could be directly used to synthesize separate NSs into a collective one. Although NSs and PFSs both have three kinds of membership functions, their property regarding membership functions is different from each other. This means that the neutrosophic aggregation operators cannot be directly used to aggregate PFSs. For example, let  $A_1 = (0, 1, 0)$  and  $A_2 = (0, 0, 1)$  be two PFSs. According to the generalized union of SVNSs defined in [32], we have  $A_1 \otimes A_2 = (0, 1, 1)$ . This result does not satisfy the condition required by PFV that the sum of its membership degrees is less than or equal to 1.

The algebraic product and algebraic sum are usually used to develop the aggregation operators for FSs [42], IFSs [43], and NSs [44, 45]. There are also other operational rules that can be used in this respect. For example, Einstein t-norms and Einstein t-conorms are two typical classes of strict Archimedean t-norms and t-conorms for aggregating a collection of intuitionistic fuzzy values (IFVs). In this direction, Wang and Liu [46] introduced the intuitionistic fuzzy aggregation operators by using Einstein operations and developed some intuitionistic fuzzy Einstein aggregation operators. Wang and

Liu [47] furthermore developed some new intuitionistic fuzzy geometric Einstein aggregation operators. Recently, Einstein t-norms and t-conorms have also been used to aggregate neutrosophic sets and neutrosophic hesitant fuzzy sets, such as the neutrosophic number-generalized weighted power averaging operator [48], and the interval neutrosophic hesitant fuzzy generalized weighted average operator [49]. Among these Einstein aggregation operators; however, some only weight the fuzzy values and others only weight the ordered positions of the fuzzy values similar to the ordered weighted averaging (OWA) operator. To overcome this drawback, Zhao and Wei [50] proposed some new intuitionistic fuzzy Einstein hybrid aggregation operators to aggregate IFVs by combining the weighted average and the OWA operator, such as the intuitionistic fuzzy Einstein hybrid averaging operator and intuitionistic fuzzy Einstein hybrid geometric averaging operator. They weight not only the given arguments but also their ordered positions. However, all of them do not satisfy the desired properties for aggregation operators such as boundedness and idempotency. Recently, some new picture fuzzy aggregation operators are proposed to apply to multicriteria group decision making, such as picture fuzzy Einstein aggregation operators, picture fuzzy-weighted average operators, picture fuzzy-weighted geometric operators, and picture fuzzy Dombi aggregation operators, etc. Although the approaches of MCGMD based on fuzzy aggregation operators have been widely investigated, the existing aggregation operators have the following shortcomings in the aggregation process:

- (1) IFS cannot well describe inconsistent, hesitation, and indeterminate information, and it does not address the influence of hesitation or neutral fuzzy information on aggregating results in the aggregation process. Thus, the MCGMD method based on intuitionistic fuzzy aggregation operators has the drawback that it may lead to unreasonable decision-making in some situations.
- (2) The neutrosophic fuzzy hybrid aggregation operators have no occasion to prove that they must satisfy the probabilistic property of the tri-membership function. In some situations, if hybrid aggregation operators are applied in specific NSs, such as intuitionistic neutrosophic sets (INSs) whose sum of tri-membership degrees is required to be less than to 1, the aggregating set is no longer an INS. For example, let  $A = (0, 0.9, 0)$  and  $B = (0, 0, 0.9)$  be two INSs. According to the inner product operation defined in [51], we obtain  $A \otimes B = (0, 0.9, 0.9)$ , which implies that the aggregating set is no longer an INS. Thus, it is important to develop new neutrosophic fuzzy hybrid aggregation operators which are suitable to all types of NSs.
- (3) Picture fuzzy multicriteria group decision making problem will be a new direction for group decision making. However, the existing Einstein aggregation operators for NSs cannot be directly applied in picture fuzzy environments. Moreover, most of these

existing neutrosophic Einstein hybrid aggregation operators do not satisfy some properties such as boundedness and idempotency for picture fuzzy sets. It is, therefore, necessary to extend the existing neutrosophic Einstein aggregation operators to picture fuzzy environments and to propose some new Einstein aggregation operators for aggregating picture fuzzy information. To the best of our knowledge, however, there are no researches on the combination between the PFSs and MCMGD, and how to aggregate PFSs is still an open problem, which is the focus of this paper.

According to the above discussions, we can see that it is better to consider both the operations rules and aggregation operators for PFSs. The main contribution of this work is as follows:

- (1) We present some picture fuzzy Einstein operational laws based on Einstein t-norms and t-conorms and discuss their desirable properties.
- (2) We introduce a new picture fuzzy Einstein hybrid-weighted aggregation (PFIEHWA) operator to aggregate PFSs. Based on the proposed aggregation operator, we develop a MAGDM method and validate its effectiveness.

The remaining sections of this paper is organized as follows. In the next section, we introduce some basic concepts related to PFSs and Einstein operations. In Section 3, by extending the Einstein t-conorm and t-norm, we develop several new Einstein operations laws for PFVs, such as generalized intersection and union, and then we discuss their desirable properties. In Section 4, we develop a picture fuzzy Einstein hybrid aggregation operator for PFVs and discuss their desirable properties. In Section 5, we apply the picture fuzzy Einstein hybrid-weighted aggregation operator to a MCGMD problem. Some numerical examples are given to verify the developed approach and to demonstrate its practicality and effectiveness. Section 6 concludes the paper.

## 2. Preliminaries

In this section, we present some basic definitions and results for IFS and PFS.

*Definition 1 [18].* An intuitionistic fuzzy set (IFS)  $A$  in a finite set  $X$  can be written as follows:

$$A = \{ \langle x, \mu_A(x), \nu_A(x), \gamma_A(x) \rangle, x \in X \}, \quad (1)$$

where  $\mu_A(x)$  and  $\nu_A(x)$ :  $X \rightarrow [0, 1]$  are, respectively, the degrees of membership and non-membership such that  $0 \leq \mu_A(x) + \nu_A(x) \leq 1$ . For each IFS  $A$  in a finite set  $X$ , the hesitancy degree of an IFS  $A$  can be expressed as  $\pi_A(x) = 1 - \mu_A(x) - \nu_A(x)$ ,  $x \in X$ . Also, we have  $0 \leq \pi_A(x) \leq 1$ , for all  $x \in X$ .

*Definition 2 [5].* A picture fuzzy set (PFS)  $A$  in a finite set  $X$  is defined as follows:

$$A = \{ \langle x, \mu_A(x), \nu_A(x), \gamma_A(x) \rangle, x \in X \}, \quad (2)$$

where  $\mu_A(x)$ ,  $\nu_A(x)$ , and  $\gamma_A(x)$  represent the positive-membership function, negative-membership function, and neutral-membership function of  $x$  to set  $A$ , respectively. For each element  $x$  in  $X$ , we have  $\mu_A(x), \nu_A(x), \gamma_A(x) \rightarrow [0, 1]$  and  $0 \leq \mu_A(x) + \nu_A(x) + \gamma_A(x) \leq 1$ .

Similar to the IFS,  $\pi_A(x) = 1 - (\mu_A(x) + \nu_A(x) + \gamma_A(x))$  could be called the refusal-membership degree of  $x$  in  $A$ . For convenience, we can use  $x = (\mu_A, \nu_A, \gamma_A)$  to represent an element in PFSs.

*Definition 3 [52].* Let  $\alpha = (\mu_\alpha, \nu_\alpha, \gamma_\alpha)$  be a PFV, and its score function  $S_\alpha$  and accuracy function  $V_\alpha$  are, respectively, defined as follows:

$$\begin{aligned} S_\alpha &= \mu_\alpha - \nu_\alpha, \\ V_\alpha &= \mu_\alpha + \nu_\alpha + \gamma_\alpha. \end{aligned} \quad (3)$$

**Theorem 1 [53].** Let  $\alpha_1$  and  $\alpha_2$  be two PFVs, and the ranking rules between them are given as follows:

- (1) If  $S(\alpha_1) > S(\alpha_2)$ , then  $\alpha_1 > \alpha_2$
- (2) If  $S(\alpha_1) < S(\alpha_2)$ , then  $\alpha_1 < \alpha_2$
- (3) If  $S(\alpha_1) = S(\alpha_2)$ , then
  - (1) If  $V(\alpha_1) = V(\alpha_2)$ , then  $\alpha_1 = \alpha_2$
  - (2) If  $V(\alpha_1) > V(\alpha_2)$ , then  $\alpha_1 > \alpha_2$

*Definition 4.* Let  $PFS(X)$  denote the set of all the PFSs in a finite set  $X$ . Given any two PFSs  $A$  and  $B$ , their inclusion, union, intersection, and complement are defined as follows:

- (1)  $A \subseteq B$  if  $\forall x \in X$ ,  $\mu_A(x) \leq \mu_B(x)$ ,  $\nu_A(x) \geq \nu_B(x)$ ,  $\gamma_A(x) \geq \gamma_B(x)$
- (2)  $A = B$  if  $\forall x \in X$ ,  $A \subseteq B$  and  $A \supseteq B$
- (3)  $A \cup B = \{x, (\max(\mu_A(x), \mu_B(x)), \min(\nu_A(x), \nu_B(x)), \min(\gamma_A(x), \gamma_B(x)))\}$
- (4)  $A \cap B = \{x, (\min(\mu_A(x), \mu_B(x)), \max(\nu_A(x), \nu_B(x)), \max(\gamma_A(x), \gamma_B(x)))\}$
- (5)  $coA = A^c = \{ \langle x, \nu_A(x), \mu_A(x), \gamma_A(x) \rangle, x \in X \}$

## 3. Picture Fuzzy Einstein Operational Laws

Triangular t-norms and t-conorms play a prominent role for aggregating fuzzy sets in group decision making. Roychowdhury and Wang [53] gave some definitions and conditions for the triangular t-norm and t-conorm, which satisfy the requirements of both conjunction and disjunction operators. The set-theoretical properties of these operators for IFSs generally hold for IFS. In the following section, triangular t-norms and t-conorms are defined as given below:

*Definition 5 [54].* A function  $T: [0, 1]^2 \rightarrow [0, 1]$  is called a t-norm if it satisfies the following four conditions:

- (1)  $T(1, x) = x, \forall x \in [0, 1]$

- (2)  $T(x, y) = T(y, x), \forall (x, y) \in [0, 1]^2$
- (3)  $T(x, T(y, z)) = T(T(x, y), z), \forall (x, y, z) \in [0, 1]^3$
- (4) If  $x \leq x'$  and  $y \leq y'$ , then  $T(x, y) \leq T(x', y')$ ,  $\forall (x, y, x', y') \in [0, 1]^4$

*Definition 6* [54]. A function  $S: [0, 1]^2 \rightarrow [0, 1]$  is called as a t-conorm if it satisfies the following four conditions:

- (1)  $S(x, 0) = 0, \forall x \in [0, 1]$
- (2)  $S(x, y) = S(y, x), \forall (x, y) \in [0, 1]^2$
- (3)  $S(x, S(y, z)) = S(S(x, y), z), \forall (x, y, z) \in [0, 1]^3$
- (4) If  $x \leq x'$  and  $y \leq y'$ , then  $S(x, y) \leq S(x', y')$ ,  $\forall (x, y, x', y') \in [0, 1]^4$

Analogous operators on fuzzy sets have also been defined on IFs. For example, the inclusion of two IFs can be defined by using the algebraic t-norm for their membership degrees and the algebraic t-conorm for their non-membership degrees, and their inclusion is still an intuitionistic fuzzy set. This is because they are related by the De Morgan duality, *i.e.*, the t-conorm  $S$  can be defined as  $S(x, y) = 1 - T(1 - x, 1 - y), \forall (x, y) \in [0, 1]^2$ . In recent years, triangular t-norms and t-conorms have also been used to define the operation laws for NSs [55]. Although PFV is a generalization of NS, their properties regarding membership functions are different from each other. This means that the operational laws for NSs cannot be directly used to aggregate PFVs. For example, with algebraic t-norm and t-conorm, the union operation for NSs is defined as  $A \cup B = (T_1 T_2, I_1 + I_2 - I_1 I_2, F_1 + F_2 - F_1 F_2)$ . However, if they are extended to PFVs, there are some limitations. For instance, let  $A = (x, 1, 0, 0)$  and  $B = (x, 0, 1, 0)$  be two PFVs. It is clear that  $B$  is the smallest one among all PFVs. According to the generalized union operation mentioned above, we obtain  $A \cup_{S,T} B = (x, 0, 1, 0)$ . This means that the aggregating result is also the smallest PFV. Therefore, the

operation “union” cannot be accepted because it is against our intuition. Furthermore, let  $C = (x, 0, 1, 0)$  and  $D = (x, 0, 0, 1)$  be two PFVs, we also have  $\mu_{C \cup D} + \nu_{C \cup D} + \gamma_{C \cup D} = 2 > 1$ , and this result does not satisfy the condition required by a PFV that the sum of its membership degrees is less than or equal to 1. Thus, it is important to develop the operational rules for PFVs. Motivated by Definition 5 and Definition 6, we first propose a generalized intersection and union for PFVs based on triangular t-norm and t-conorm.

*Definition 7.* Let  $\alpha_j = (\mu_j, \nu_j, \gamma_j), (j = 1, 2)$  be two PFVs. The generalized intersection and union between  $\alpha_1$  and  $\alpha_2$  are, respectively, defined as follows:

$$\begin{aligned} \alpha_1 \oplus_{S,T} \alpha_2 &= (S(\mu_1, \mu_2), T(\nu_1, \nu_2), T(\gamma_1, \gamma_2)), \\ \alpha_1 \otimes_{S,T} \alpha_2 &= (T(\mu_1, \mu_2), S(\nu_1, \nu_2), T(\gamma_1, \gamma_2)). \end{aligned} \quad (4)$$

There are lots of operators based on the algebraic operation which are one of the general concepts of the t-norms and t-conorms. Einstein operation, including the Einstein product and the Einstein sum, also belongs to the t-norms and t-conorms families. Let additive generator  $N(x) = 1 - x$ , then Einstein product  $\otimes_\varepsilon$  and Einstein sum  $\oplus_\varepsilon$  are defined as follows [36]:

- (1)  $\alpha \otimes_\varepsilon \beta = \alpha\beta / (1 + (1 - \alpha)(1 - \beta)), \forall (\alpha, \beta) \in [0, 1]$
- (2)  $\alpha \oplus_\varepsilon \beta = (\alpha + \beta) / (1 + \alpha\beta), \forall (\alpha, \beta) \in [0, 1]$

With the abovementioned analysis, the operations laws for PFVs based on Einstein t-norms and t-conorms can be defined as follows:

*Definition 8.* Let  $\alpha = (\mu, \nu, \gamma)$  and  $\alpha_j = (\mu_j, \nu_j, \gamma_j), j = 1, 2$ , be PFVs, and  $\lambda$  is a positive real number. Let additive generator  $N(x) = 1 - x$ , then we have the following operations laws:

$$\alpha_1 \oplus_\varepsilon \alpha_2 = \left( \frac{\mu_1 + \mu_2}{1 + \mu_1 \mu_2}, \frac{\nu_1 \nu_2}{1 + (1 - \nu_1)(1 - \nu_2)}, \frac{\gamma_1 \gamma_2}{1 + (1 - \gamma_1)(1 - \gamma_2)} \right), \quad (5)$$

$$\alpha_1 \otimes_\varepsilon \alpha_2 = \left( \frac{\mu_1 \mu_2}{1 + (1 - \mu_1)(1 - \mu_2)}, \frac{\nu_1 + \nu_2}{1 + \nu_1 \nu_2}, \frac{\gamma_1 \gamma_2}{1 + (1 - \gamma_1)(1 - \gamma_2)} \right), \quad (6)$$

$$\lambda \cdot_\varepsilon \alpha = \left( \frac{(1 + \mu)^\lambda - (1 - \mu)^\lambda}{(1 + \mu)^\lambda + (1 - \mu)^\lambda}, \frac{2\nu^\lambda}{(2 - \nu)^\lambda + \nu^\lambda}, \frac{2\gamma^\lambda}{(2 - \gamma)^\lambda + \gamma^\lambda} \right), \quad (7)$$

$$\alpha^\lambda = \left( \frac{2\mu^\lambda}{(2 - \mu)^\lambda + \mu^\lambda}, \frac{(1 + \nu)^\lambda - (1 - \nu)^\lambda}{(1 + \nu)^\lambda + (1 - \nu)^\lambda}, \frac{2\gamma^\lambda}{(2 - \gamma)^\lambda + \gamma^\lambda} \right). \quad (8)$$

**Theorem 2.** Let  $\alpha_j = (\mu_j, \nu_j, \gamma_j), j = 1, 2$ , and  $\alpha = (\mu, \nu, \gamma)$  be PFVs,  $\lambda$  is a positive real number. Then,  $\alpha_1 \oplus_\varepsilon \alpha_2, \alpha_1 \otimes_\varepsilon \alpha_2, \lambda \cdot_\varepsilon \alpha$ , and  $\alpha^\lambda$  are also a PFV.

*Proof:* See Appendix A. □

**Theorem 3.** Let  $\lambda_j \in [0, 1]$ ,  $i = 1, 2, 3$ , and  $\alpha = (\mu, \nu, \gamma)$  be PFVs, and  $\lambda_1, \lambda_2$  and  $\lambda$  are positive real numbers, then we have the following properties:

$$\alpha_1 \oplus_\varepsilon \alpha_2 = \alpha_2 \oplus_\varepsilon \alpha_1, \quad (9)$$

$$\alpha_1 \otimes_\varepsilon \alpha_2 = \alpha_2 \otimes_\varepsilon \alpha_1, \quad (10)$$

$$(\alpha_1 \oplus_\varepsilon \alpha_2) \oplus_\varepsilon \alpha_3 = \alpha_1 \oplus_\varepsilon (\alpha_2 \oplus_\varepsilon \alpha_3), \quad (11)$$

$$\lambda(\alpha_1 \oplus_\varepsilon \alpha_2) = \lambda \alpha_1 \oplus_\varepsilon \lambda \alpha_2, \quad (12)$$

$$(\alpha_1 \otimes_\varepsilon \alpha_2)^\lambda = \alpha_1^\lambda \otimes_\varepsilon \alpha_2^\lambda, \quad (13)$$

$$\lambda_1 \cdot_\varepsilon \alpha \oplus_\varepsilon \lambda_2 \cdot_\varepsilon \alpha = (\lambda_1 + \lambda_2) \cdot_\varepsilon \alpha, \quad (14)$$

$$\alpha^{\lambda_1} \otimes_\varepsilon \alpha^{\lambda_2} = \alpha^{(\lambda_1 + \lambda_2)}. \quad (15)$$

Note that the proofs of these theorems are straightforward and thus omitted here for the sake of brevity.

In the next section, we investigate an Einstein hybrid aggregation operator under the picture fuzzy environment based on Einstein operations.

#### 4. Picture Fuzzy Einstein Hybrid-Weighted Aggregation Operator

In this section, we propose the picture fuzzy Einstein hybrid-weighted average (PFEHWA) operator based on the proposed Einstein operations laws on picture fuzzy values.

**Definition 9.** Let  $\alpha_j = (\mu_{\alpha_j}, \nu_{\alpha_j}, \gamma_{\alpha_j})$ , ( $j = 1, 2, \dots, n$ ) be a collection of PFVs. The picture fuzzy Einstein hybrid-weighted average (PFIEHWA) operator is a mapping  $\text{PFIEHWA}: \Omega^n \rightarrow \Omega$ , with an aggregation-associated weighting vector  $\omega = (\omega_1, \omega_2, \dots, \omega_n)^T$ ,  $\omega_j \in [0, 1]$  and  $\sum_{j=1}^n \omega_j = 1$ , such that

$$\text{PFIEHWA}_{\omega, \lambda}(A_1, A_2, \dots, A_n) = \frac{\oplus_{\varepsilon, j=1}^n (\omega_{\varepsilon(j)} \lambda_j A_j)}{\sum_{j=1}^n \omega_{\varepsilon(j)} \lambda_j}, \quad (16)$$

where  $\varepsilon: (j) \rightarrow (1, 2, \dots, n)$  is a permutation such that  $\alpha_j$  is the  $\varepsilon(j)$ <sup>th</sup> largest element of the collection of PFVs and  $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n)^T$  is the weighting vector of  $\alpha_j$ , with  $\lambda_j \in [0, 1]$  and  $\sum_{j=1}^n \lambda_j = 1$ .

**Theorem 4.** For a collection of PFVs  $\omega_j \in [0, 1]$ ,  $j = 1, 2, \dots, n$ , their aggregated value by using the PFIEHWA operator is also a PFV, and

$$\begin{aligned} & \text{PFIEHWA}_{\omega, \lambda}(\alpha_1, \alpha_2, \dots, \alpha_n) \\ &= \left( \frac{\prod_{j=1}^n (1 + \mu_{\alpha_j})^{(\omega_{\varepsilon(j)} \lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)} \lambda_j)} - \prod_{j=1}^n (1 - \mu_{\alpha_j})^{(\omega_{\varepsilon(j)} \lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)} \lambda_j)}}{\prod_{j=1}^n (1 + \mu_{\alpha_j})^{(\omega_{\varepsilon(j)} \lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)} \lambda_j)} + \prod_{j=1}^n (1 - \mu_{\alpha_j})^{(\omega_{\varepsilon(j)} \lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)} \lambda_j)}} \right. \\ & \quad \left. \frac{2 \prod_{j=1}^n (\nu_{\alpha_j})^{(\omega_{\varepsilon(j)} \lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)} \lambda_j)}}{\prod_{j=1}^n (2 - (\nu_{\alpha_j}))^{(\omega_{\varepsilon(j)} \lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)} \lambda_j)} + \prod_{j=1}^n (\nu_{\alpha_j})^{(\omega_{\varepsilon(j)} \lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)} \lambda_j)}} \right. \\ & \quad \left. \frac{2 \prod_{j=1}^n (\gamma_{\alpha_j})^{(\omega_{\varepsilon(j)} \lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)} \lambda_j)}}{\prod_{j=1}^n (2 - (\gamma_{\alpha_j}))^{(\omega_{\varepsilon(j)} \lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)} \lambda_j)} + \prod_{j=1}^n (\gamma_{\alpha_j})^{(\omega_{\varepsilon(j)} \lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)} \lambda_j)}} \right). \end{aligned} \quad (17)$$

*Proof:* See Appendix B.  $\square$

**Theorem 5.** The PFIEHWA operator satisfies the following properties:

(1) (Idempotency): If  $\alpha_j = (\mu_{\alpha_j}, \nu_{\alpha_j}, \gamma_{\alpha_j})$ ,  $j = 1, 2, \dots, n$  are all equal, that is,  $\alpha_j = \alpha$ , for all  $j = 1, 2, \dots, n$ . Then,

$$\text{PFIEHWA}_{\omega, \lambda}(\alpha_1, \alpha_2, \dots, \alpha_n) = \alpha. \quad (18)$$

(2) (Boundedness): Let  $\alpha_{\min} = (\min_{1 \leq j \leq n} \mu_{\alpha_j}, \max_{1 \leq j \leq n} \nu_{\alpha_j}, \max_{1 \leq j \leq n} \gamma_{\alpha_j})$  and  $\alpha_{\max} = (\max_{1 \leq j \leq n} \mu_{\alpha_j}, \min_{1 \leq j \leq n} \nu_{\alpha_j}, \min_{1 \leq j \leq n} \gamma_{\alpha_j})$ , then

$$\alpha_{\min} \leq \text{PFIEHWA}_{\omega, \lambda}(\alpha_1, \alpha_2, \dots, \alpha_n) \leq \alpha_{\max}. \quad (19)$$

(3) (Monotonicity): Let  $\alpha_j^1 = (\mu_{\alpha_j^1}, \nu_{\alpha_j^1}, \gamma_{\alpha_j^1})$ ,  $j = 1, 2, \dots, n$  and  $\alpha_j^2 = (\mu_{\alpha_j^2}, \nu_{\alpha_j^2}, \gamma_{\alpha_j^2})$ ,  $j = 1, 2, \dots, n$ , be two collections of PFVs, if  $\mu_{\alpha_j^1} \leq \mu_{\alpha_j^2}$ ,  $\nu_{\alpha_j^1} \geq \nu_{\alpha_j^2}$ , and  $\gamma_{\alpha_j^1} \geq \gamma_{\alpha_j^2}$ . Then,

$$\text{PFIEHWA}_{\omega, \lambda}(\alpha_1^1, \alpha_2^1, \dots, \alpha_n^1) \leq \text{PFIEHWA}_{\omega, \lambda}(\alpha_1^2, \alpha_2^2, \dots, \alpha_n^2). \quad (20)$$

*Proof:* See Appendix C.  $\square$

## 5. Application of Picture Einstein Fuzzy Hybrid Aggregation Operator

In this section, we investigate the application of the picture Einstein fuzzy hybrid aggregation operator in group decision making with picture fuzzy information.

### 5.1. MAGDM Method Based on the PFIEHWA Operator.

For a MCGMD problem under the picture fuzzy environment, let  $D = (d_1, d_2, \dots, d_l)$  be a set of decision makers,  $C = (C_1, C_2, \dots, C_n)$  be a set of criteria, and  $X = (X_1, X_2, \dots, X_m)$  be a set of alternatives to be evaluated. Let  $\tilde{D}^{(k)} = (\tilde{\alpha}_{ij}^{(k)})_{m \times n}$  be a picture fuzzy decision matrix, where  $\tilde{\alpha}_{ij}^{(k)} = (\mu_{\alpha_{ij}}^{(k)}, \nu_{\alpha_{ij}}^{(k)}, \gamma_{\alpha_{ij}}^{(k)})$  is a PFV for alternative  $X_i$  with respect to criteria  $C_j$  provided by decision maker  $d_k$ , such that  $0 \leq \mu_{\alpha_{ij}}^{(k)} \leq 1$ ,  $0 \leq \nu_{\alpha_{ij}}^{(k)} \leq 1$ ,  $0 \leq \gamma_{\alpha_{ij}}^{(k)} \leq 1$ ,  $0 \leq \mu_{\alpha_{ij}}^{(k)} + \nu_{\alpha_{ij}}^{(k)} + \gamma_{\alpha_{ij}}^{(k)} \leq 1$ ,  $k = 1, 2, \dots, l$ ,  $i = 1, 2, \dots, m$  and  $j = 1, 2, \dots, n$ . The expert committee gives the weighting vector  $\lambda = (\lambda^{(1)}, \lambda^{(2)}, \dots, \lambda^{(l)})^T$  for the decision makers, where  $\lambda^{(k)} \in [0, 1]$  and  $\sum_{k=1}^l \lambda^{(k)} = 1$ . Considering that different decision makers are familiar with differentiated fields, the expert committee also determines the ordering weights vector  $\omega = (\omega^{(1)}, \omega^{(2)}, \dots, \omega^{(l)})^T$  for the decision makers, where  $\omega^{(k)} \in [0, 1]$  and  $\sum_{k=1}^l \omega^{(k)} = 1$ . After that, the picture fuzzy decision matrixes  $\tilde{D}^{(k)} = (\tilde{\alpha}_{ij}^{(k)})_{m \times n}$  will be aggregated into a collective picture fuzzy decision matrix  $\tilde{D} = (\tilde{\alpha}_{ij})_{m \times n}$ . Then, the expert committee assigns the weighting vector  $\eta = (\eta_1, \eta_2, \dots, \eta_n)^T$  for the criteria according to their relative importance in decision making, where  $\eta_j \in [0, 1]$  and  $\sum_{j=1}^n \eta_j = 1$ . Meanwhile, considering the fact that diverse alternatives may have differentiated focuses and advantages, the expert committee also gives the aggregation-associated weights vector  $\xi = (\xi_1, \xi_2, \dots, \xi_n)^T$  for different criteria, where  $\xi_j \in [0, 1]$  and  $\sum_{j=1}^n \xi_j = 1$ .

The complete procedure for multicriteria group decision making based on the proposed picture fuzzy Einstein hybrid-weighted aggregation operator can be summarized as follows:

*Step 1.* In order to eliminate the impact of different types of criteria values (i.e., benefit criteria or cost criteria), we will transform the criteria values of cost type into those of the benefit type i.e., transform  $\tilde{D}^{(k)} = (\tilde{\alpha}_{ij}^{(k)})_{m \times n}$  into a normalized picture fuzzy decision matrix  $\tilde{R}^{(k)} = (\tilde{r}_{ij}^{(k)})_{m \times n}$ , where

$$\tilde{r}_{ij}^{(k)} = \begin{cases} \alpha_{ij}^k & \text{for benefit criteria} \\ (\alpha_{ij}^k)^c & \text{for cost criteria} \end{cases}, \quad (21)$$

where  $(\alpha_{ij}^k)^c$  is the complement of  $\alpha_{ij}^k$  such that  $(\alpha_{ij}^k)^c = (\nu_{ij}^{(k)}, \mu_{ij}^{(k)}, \gamma_{ij}^{(k)})$ .

*Step 2.* Utilizing the PFIEHWA operator to aggregate picture fuzzy decision matrixes  $\tilde{R}^{(k)} = (\tilde{r}_{ij}^{(k)})_{m \times n}$  into a collective picture fuzzy decision matrix  $\tilde{R} = (\tilde{r}_i^{(k)})_{m \times k}$ .

$$\tilde{r}_i^{(k)} = \text{PFIEHWA}_{\omega, \lambda}(\tilde{r}_{i1}^{(k)}, \tilde{r}_{i2}^{(k)}, \dots, \tilde{r}_{in}^{(k)}) = \oplus_{\varepsilon=j=1}^n \frac{(\xi_{\varepsilon(ij)}^{(k)} \lambda_j \tilde{r}_{ij}^{(k)})}{\sum_{k=1}^l \xi_{\varepsilon(ij)}^{(k)} \lambda_j}. \quad (22)$$

*Step 3.* Utilizing the PFIEHWA  $\tilde{r}_i$  to aggregate all the evaluation values  $\tilde{r}_i^{(k)}$  into a collective evaluation value  $\tilde{r}_i$  for alternative  $X_i$ .

$$\tilde{r}_i = \text{PFIEHWA}_{\xi, \eta}(\tilde{r}_i^{(1)}, \tilde{r}_i^{(2)}, \tilde{r}_i^{(3)}, \dots, \tilde{r}_i^{(l)}) = \oplus_{\varepsilon=k=1}^l \frac{(\omega_{\varepsilon(ik)} \eta_k \tilde{r}_i^{(k)})}{\sum_{k=1}^l \omega_{\varepsilon(ik)} \eta_k}. \quad (23)$$

*Step 4.* Ranking  $\tilde{r}_i$  by using the ranking method described in Section 2 and select the best one.

*5.2. Illustrative Example.* In this section, we use an example presented in [49] to illustrate the proposed method.

*Example 1 [49].* Suppose a company wants to develop a new career, where there are three alternatives:  $X_1, X_2$ , and  $X_3$  to be selected.  $X_1$  is the real estate industry,  $X_2$  is the food industry, and  $X_3$  is the education industry. After preliminary screening, three experts  $d_1, d_2$ , and  $d_3$  are asked to evaluate the alternatives. Four criteria are determined: the ability to compete ( $C_1$ ), the ability to grow ( $C_2$ ), the influence of the surrounding environment ( $C_3$ ), and the influence of social politics ( $C_4$ ). Assume that the subjective importance degree of each decision maker is  $\lambda = (0.2, 0.4, 0.4)^T$ . Simultaneously, considering that some decision makers could be more familiar with career management, the aggregation-associated weighting vector of decision maker  $d_k$  is  $\omega = (0.3, 0.3, 0.4)^T$ . Furthermore, suppose the weight vector and aggregation-associated weight vector for the four criteria are  $\eta = (0.2, 0.3, 0.3, 0.2)^T$  and  $\xi = (0.2, 0.1, 0.3, 0.4)^T$ , respectively. The picture fuzzy decision matrices are given as follows:

$$\begin{aligned}
 \tilde{D}^{(1)} &= \begin{bmatrix} (0.1, 0.3, 0.5) & (0.5, 0.2, 0.3) & (0.2, 0.2, 0.1) & (0.3, 0.1, 0.2) \\ (0.2, 0.2, 0.3) & (0.1, 0.2, 0.4) & (0.2, 0.1, 0.3) & (0.1, 0.2, 0.3) \\ (0.1, 0.3, 0.5) & (0.3, 0.1, 0.2) & (0.3, 0.2, 0.1) & (0.2, 0.3, 0.2) \end{bmatrix}, \\
 \tilde{D}^{(2)} &= \begin{bmatrix} (0.3, 0.1, 0.2) & (0.3, 0.1, 0.2) & (0.3, 0.1, 0.2) & (0.2, 0.2, 0.3) \\ (0.1, 0.3, 0.2) & (0.2, 0.1, 0.3) & (0.3, 0.3, 0.2) & (0.3, 0.1, 0.1) \\ (0.2, 0.3, 0.1) & (0.2, 0.1, 0.3) & (0.4, 0.1, 0.3) & (0.2, 0.1, 0.3) \end{bmatrix}, \\
 \tilde{D}^{(3)} &= \begin{bmatrix} (0.5, 0.1, 0.1) & (0.5, 0.2, 0.1) & (0.2, 0.4, 0.1) & (0.3, 0.1, 0.1) \\ (0.4, 0.1, 0.2) & (0.3, 0.2, 0.4) & (0.2, 0.1, 0.4) & (0.5, 0.2, 0.1) \\ (0.4, 0.2, 0.2) & (0.4, 0.2, 0.1) & (0.5, 0.2, 0.3) & (0.3, 0.2, 0.3) \end{bmatrix}.
 \end{aligned} \tag{24}$$

Step 1. Considering the criteria are all the benefit criteria, there is no need to transform them into benefits ones. Thus we have  $\tilde{D}^{(k)} = \tilde{R}^{(k)}$

Step 2. Utilizing the PFIEHWA operator to aggregate all individual picture fuzzy decision matrix  $\tilde{R}^{(k)}$  into a collective picture fuzzy decision matrix  $\tilde{R}$

$$(1 - \mu_{A_j}^1) / (1 - \mu_{A_j}^1) \leq (1 + \mu_{A_j}^2) / (1 + \mu_{A_j}^2)$$

Let us illustrate this step by using  $\tilde{r}_{11}$  as an example. Since  $\tilde{r}_{11}^{(1)} = (0.1, 0.3, 0.5)$ ,  $\tilde{r}_{12}^{(1)} = (0.5, 0.2, 0.3)$ ,  $\tilde{r}_{13}^{(1)} = (0.2, 0.2, 0.1)$ ,  $\tilde{r}_{14}^{(4)} = (0.3, 0.2, 0.1)$ ,  $\eta = (0.2, 0.3, 0.3, 0.2)^T$ , and  $\xi = (0.2, 0.1, 0.3, 0.4)^T$ , we have  $S(\tilde{r}_{11}^{(1)}) = -0.2$ ,  $S(\tilde{r}_{12}^{(1)}) = 0.3$ ,  $S(\tilde{r}_{13}^{(1)}) = 0$ , and  $S(\tilde{r}_{14}^{(1)}) = 0.1$  by using the ranking function given in Theorem 1. Thus, we can obtain  $S(\tilde{r}_{11}^{(1)}) < S(\tilde{r}_{13}^{(1)}) < S(\tilde{r}_{14}^{(1)}) < S(\tilde{r}_{12}^{(1)})$ , which implies that  $\tilde{r}_{11}^{(1)} < \tilde{r}_{13}^{(1)} < \tilde{r}_{14}^{(1)} < \tilde{r}_{12}^{(1)}$ . Hence, we have  $\varepsilon_{11}(1) = 4$ ,  $\varepsilon_{12}(1) = 3$ ,  $\varepsilon_{13}(1) = 2$ , and  $\varepsilon_{14}(1) = 1$ . Then,

$$\begin{aligned}
 \frac{\xi_{\varepsilon_{11}(1)} \eta_1^{(1)}}{\sum_{j=1}^4 \xi_{\varepsilon_{1j}} \eta_1^{(1)}} &= \frac{0.2 \times 0.4}{0.2 \times 0.4 + 0.3 \times 0.3 + 0.3 \times 0.1 + 0.2 \times 0.2} = 0.3333, \\
 \frac{\xi_{\varepsilon_{12}(1)} \eta_2^{(1)}}{\sum_{j=1}^4 \xi_{\varepsilon_{1j}} \eta_2^{(1)}} &= \frac{0.3 \times 0.3}{0.2 \times 0.4 + 0.3 \times 0.3 + 0.3 \times 0.1 + 0.2 \times 0.2} = 0.3750, \\
 \frac{\xi_{\varepsilon_{13}(1)} \eta_3^{(1)}}{\sum_{j=1}^4 \xi_{\varepsilon_{1j}} \eta_3^{(1)}} &= \frac{0.3 \times 0.1}{0.2 \times 0.4 + 0.3 \times 0.3 + 0.3 \times 0.1 + 0.2 \times 0.2} = 0.1250, \\
 \frac{\xi_{\varepsilon_{14}(1)} \eta_4^{(1)}}{\sum_{j=1}^4 \xi_{\varepsilon_{1j}} \eta_4^{(1)}} &= \frac{0.2 \times 0.2}{0.2 \times 0.4 + 0.3 \times 0.3 + 0.3 \times 0.1 + 0.2 \times 0.2} = 0.1667.
 \end{aligned} \tag{25}$$

Finally, we have  $PFIEHWA_{\xi, \eta}(\tilde{r}_{11}^{(1)}, \tilde{r}_{12}^{(1)}, \tilde{r}_{13}^{(1)}, \tilde{r}_{14}^{(1)}) = (0.1720, 0.2054, 0.2958)$  by using (21). Similarly, we can derive the collective evaluation matrix given in Table 1.

Step 3. Utilizing the PFIEHWA operator to aggregate all  $\tilde{r}_i^{(k)}$ ,  $k = 1, 2, 3$ , into a collective PFV  $\tilde{r}_i$  for alternative  $X_i$  over all the experts. According to Theorem 1, we have

TABLE 1: Collective picture fuzzy decision matrix.

	$D_1$	$D_2$	$D_3$
$X_1$	(0.1720, 0.2054, 0.2958)	(0.1355, 0.1265, 0.2296)	(0.1728, 0.1879, 0.100)
$X_2$	(0.0700, 0.1700, 0.3334)	(0.1057, 0.1679, 0.2090)	(0.1655, 0.1219, 0.2764)
$X_3$	(0.1109, 0.1485, 0.2575)	(0.1174, 0.1210, 0.2522)	(0.2007, 0.2000, 0.1753)

$S(\tilde{r}_1^{(1)}) = -0.0334$ ,  $S(\tilde{r}_1^{(2)}) = -0.1$  and  $S(\tilde{r}_1^{(3)}) = -0.0376$ . Thus,  $S(\tilde{r}_1^{(3)}) < S(\tilde{r}_1^{(1)}) < S(\tilde{r}_1^{(2)})$ , which means that  $\tilde{r}_1^{(3)} < \tilde{r}_1^{(1)} < \tilde{r}_1^{(2)}$ . Hence,  $\delta_{11} = 2$ ,  $\delta_{12} = 3$  and  $\delta_{13} = 1$ . Then, we have as follows:

$$\begin{aligned}\omega_{\varepsilon(11)} \frac{\lambda_1}{\sum_{k=1}^3 \omega_{\varepsilon(1k)} \lambda_k} &= \frac{(0.2 \times 0.3)}{(0.2 \times 0.3 + 0.4 \times 0.3 + 0.4 \times 0.4)} = 0.1765, \\ \omega_{\varepsilon(12)} \frac{\lambda_2}{\sum_{k=1}^3 \omega_{\varepsilon(1k)} \lambda_k} &= \frac{(0.4 \times 0.3)}{(0.2 \times 0.3 + 0.4 \times 0.3 + 0.4 \times 0.4)} = 0.3529, \\ \omega_{\varepsilon(13)} \frac{\lambda_3}{\sum_{k=1}^3 \omega_{\varepsilon(1k)} \lambda_k} &= \frac{(0.4 \times 0.4)}{(0.2 \times 0.3 + 0.4 \times 0.3 + 0.4 \times 0.4)} = 0.4706.\end{aligned}\quad (26)$$

Thus, we can derive the following result:

$$\tilde{r}_1 = \text{PFIEHWA}_{\omega, \lambda}(\tilde{r}_1^{(1)}, \tilde{r}_2^{(1)}, \tilde{r}_3^{(1)}, \tilde{r}_4^{(1)}) = (0.1129, 0.1650, 0.2894). \quad (27)$$

Similarly, we can obtain the following results:

$$\begin{aligned}\tilde{r}_2 &= \text{PFIEHWA}_{\omega, \lambda}(\tilde{r}_1^{(2)}, \tilde{r}_2^{(2)}, \tilde{r}_3^{(2)}, \tilde{r}_4^{(2)}) = (0.1156, 0.1424, 0.2271), \\ \tilde{r}_3 &= \text{PFIEHWA}_{\omega, \lambda}(\tilde{r}_1^{(3)}, \tilde{r}_2^{(3)}, \tilde{r}_3^{(3)}, \tilde{r}_4^{(3)}) = (0.1812, 0.1639, 0.1821).\end{aligned}\quad (28)$$

*Step 4.* Computing the ranking values  $R(\tilde{r}_i)$  for  $A_i$ , for all  $k=1, 2$ , and  $3$ . By using Theorem 1, we have  $S(\tilde{r}_1) = -0.0521$ ,  $S(\tilde{r}_2) = -0.0268$ , and  $S(\tilde{r}_3) = 0.0173$ . Since  $S(\tilde{r}_1) < S(\tilde{r}_2) < S(\tilde{r}_3)$ , we obtain  $X_1 < X_2 < X_3$ , which implies that  $X_3$  is the most desirable career alternative. This order is the same as the method proposed by Liu and Shi's work.

**5.3. Comparative Analysis.** In this section, we compare the proposed method based on the *PFEHWA* operator with the method given by Zhang [44], Peng [56], and Liu [48] using the following example in [44]. In the method proposed by Zhang [44], a quasi-intuitionistic fuzzy Einstein hybrid-weighted averaging (*QIFEHWA*) operator and intuitionistic fuzzy Einstein weighted averaging (*IFEWA*) operator were utilized, which does not consider hesitation or neutral information in the aggregation process, whereas the method given by Peng [56], Liu [48], and the proposed method considers hesitation or neutral information. In the method of Peng [56] and Liu [48], NS was used to develop some single-valued Neutrosophic fuzzy-weighted averaging operators, such as the Einstein single-valued neutrosophic

number-weighted averaging (*ESVNNWA*) operator in [48] and the simplified neutrosophic hybrid ordered weighted averaging (*SNNHOWA*) operator in [56].

*Example 2 [44].* Suppose a computer center in a university wants to select a new information system from the following four possible alternatives:  $X_1$ ,  $X_2$ ,  $X_3$ , and  $X_4$ . Three decision makers are asked to make a decision according to the following four criteria: (1)  $C_1$  is the costs of the hardware and software investment; (2)  $C_2$  is the contribution to organization performance; (3)  $C_3$  is the effort to transition from the current systems; and (4)  $C_4$  is the reliability of outsourcing software development; these are all benefit type criteria. Suppose the weights vector of decision makers  $d_1$ ,  $d_2$ , and  $d_3$  is  $\lambda = (\lambda^1, \lambda^2, \lambda^3)^T = (0.2, 0.5, 0.3)^T$  and the associated-weighting vector is  $\omega = (\omega^1, \omega^2, \omega^3)^T = (1/3, 1/3, 1/3)^T$ . Then, the decision maker  $d_k$  determines the weight vector of four criteria, which are  $\eta_j^{(1)} = (0.4, 0.3, 0.1, 0.2)^T$ ,  $\eta_j^{(2)} = (0.1, 0.3, 0.5, 0.1)^T$ , and  $\eta_j^{(3)} = (0.1, 0.2, 0.3, 0.4)^T$ , where  $j = 1, 2, 3, 4$ . According to their preferences, the decision makers also give the associated-weighting vectors of the four criteria:  $\xi_j^{(1)} = (0.4, 0.3, 0.2, 0.1)^T$ ,  $\xi_j^{(2)} = (0.3, 0.3, 0.2, 0.2)^T$ , and



TABLE 2: Picture fuzzy decision matrix  $\tilde{D}^{(1)}$ .

	$C_1$	$C_2$	$C_3$	$C_4$
$X_1$	(0.5, 0.4, 0.1)	(0.4, 0.3, 0.3)	(0.5, 0.3, 0.2)	(0.2, 0.6, 0.2)
$X_2$	(0.5, 0.4, 0.1)	(0.3, 0.7, 0.0)	(0.2, 0.8, 0.0)	(0.4, 0.5, 0.1)
$X_3$	(0.2, 0.6, 0.2)	(0.8, 0.1, 0.1)	(0.6, 0.4, 0.0)	(0.1, 0.7, 0.2)
$X_4$	(0.1, 0.9, 0.0)	(0.2, 0.8, 0.0)	(0.7, 0.2, 0.1)	(0.4, 0.6, 0.0)

TABLE 3: Picture fuzzy decision matrix  $\tilde{D}^{(2)}$ .

	$C_1$	$C_2$	$C_3$	$C_4$
$X_1$	(0.3, 0.6, 0.1)	(0.2, 0.7, 0.1)	(0.5, 0.5, 0.0)	(0.5, 0.3, 0.2)
$X_2$	(0.3, 0.7, 0.0)	(0.6, 0.4, 0.0)	(0.7, 0.2, 0.1)	(0.4, 0.5, 0.1)
$X_3$	(0.6, 0.3, 0.1)	(0.4, 0.4, 0.2)	(0.2, 0.7, 0.1)	(0.3, 0.6, 0.1)
$X_4$	(0.2, 0.5, 0.3)	(0.5, 0.3, 0.2)	(0.5, 0.4, 0.1)	(0.4, 0.3, 0.3)

TABLE 4: Picture fuzzy decision matrix  $\tilde{D}^{(3)}$ .

	$C_1$	$C_2$	$C_3$	$C_4$
$X_1$	(0.7, 0.3, 0.0)	(0.4, 0.5, 0.1)	(0.5, 0.4, 0.1)	(0.6, 0.2, 0.2)
$X_2$	(0.5, 0.5, 0.0)	(0.3, 0.5, 0.2)	(0.8, 0.1, 0.1)	(0.7, 0.1, 0.2)
$X_3$	(0.8, 0.2, 0.0)	(0.2, 0.3, 0.5)	(0.6, 0.3, 0.1)	(0.2, 0.7, 0.1)
$X_4$	(0.9, 0.1, 0.0)	(0.8, 0.1, 0.1)	(0.2, 0.1, 0.7)	(0.2, 0.6, 0.2)

TABLE 5: Ranking results for example 2.

Methods	Ranking values	Ranking results
Zhang method [44] (QIFEHWA and IFEWA operator)	$S(r_1) = 0.0913$ ; $S(r_2) = 0.3768$ ; $S(r_3) = 0.0736$ ; $S(r_4) = 0.1769$ .	$X_2 \succ X_4 \succ X_1 \succ X_3$
Peng method [58] (SNNHOWA operator)	$S(r_1) = 0.7609$ ; $S(r_2) = 0.8456$ ; $S(r_3) = 0.7281$ ; $S(r_4) = 0.7577$ .	$X_2 \succ X_1 \succ X_4 \succ X_3$
Liu method [48] (ESVNNWA operator)	$S(r_1) = 0.6486$ ; $S(r_2) = 0.7550$ ; $S(r_3) = 0.6519$ ; $S(r_4) = 0.6912$ .	$X_2 \succ X_4 \succ X_3 \succ X_1$
Our method (PFEHWA operator)	$S(r_1) = 0.4655$ ; $S(r_2) = 0.6142$ ; $S(r_3) = 0.5080$ ; $S(r_4) = 0.5543$ .	$X_2 \succ X_4 \succ X_3 \succ X_1$

$\xi_j^{(3)} = (0.5, 0.3, 0.1, 0.1)^T$ , where  $j = 1, 2, 3, 4$ . Considering that the value of each criterion is evaluated using the intuitionistic fuzzy information by the decision makers, we first transform IFVs into PFVs, as shown in Tables 2-4.

A comparison of the ranking results using different methods is shown in Table 5. We can see from Table 5 that different methods lead to different rankings. However,  $X_2$  is always the best alternative. We can also see that the method of Liu and our method generate the same ranking for four alternatives. The reason can be explained as follows. First, the shortcomings of IFSs and NSs that were discussed earlier can account for the differences in the final rankings. Second, these methods apply different types of ranking measures to rank IFSs or NSs, such as the score function or accuracy function. In the proposed method, a novel ranking function based on the parametric distance measure can distinguish two different PFVs while other ranking measures could not.

## 6. Conclusion

In this study, we have proposed a hybrid Einstein aggregation operator on the basis of the proposed picture Einstein operations for aggregating picture fuzzy information and

investigated their application in multicriteria group decision making. First, based on the picture Einstein operation laws, we have developed a new operator for aggregating PFVs. Then, we have utilized the proposed operator to develop a method for MCGDM problems in which the evaluation values are represented by PFVs. Finally, an example is conducted to illustrate the practicality and effectiveness of the proposed MCGDM approach. In future research, the proposed method could be extended to interval-valued picture fuzzy sets and trapezoidal picture fuzzy sets. Another interesting direction could be to develop other types of picture fuzzy aggregation operators.

## Appendix

### A. Proof of Theorem 2

*Proof:* We first prove that  $\alpha_1 \oplus_\epsilon \alpha_2$  is a PFV. Since  $0 \leq \mu_1, \mu_2, \nu_1, \nu_2, \gamma_1, \gamma_2 \leq 1$ ,  $0 \leq \mu_1 + \nu_1 + \gamma_1 \leq 1$  and  $0 \leq \mu_2 + \nu_2 + \gamma_2 \leq 1$ , we obtain  $1 - \nu_1 \geq \mu_1$ ,  $1 - \nu_2 \geq \mu_2$ ,  $1 - \gamma_1 \geq \mu_1$ , and  $1 - \gamma_2 \geq \mu_2$ . Thus, we have as follows:



$$\begin{aligned}
& \frac{\mu_1 + \mu_2}{1 + \mu_1\mu_2} + \frac{\nu_1\nu_2}{1 + (1 - \nu_1)(1 - \nu_2)} + \frac{\gamma_1\gamma_2}{1 + (1 - \gamma_1)(1 - \gamma_2)} \leq \frac{\mu_1 + \mu_2}{1 + \mu_1\mu_2} + \frac{\nu_1\nu_2}{1 + \mu_1\mu_2} + \frac{\gamma_1\gamma_2}{1 + \mu_1\mu_2} \\
& = \frac{\mu_1 + \mu_2 + \nu_1\nu_2 + \gamma_1\gamma_2}{1 + \mu_1\mu_2} \\
& \leq \frac{\mu_1 + \mu_2 + (1 - (\mu_1 + \gamma_1))(1 - (\mu_2 + \gamma_2)) + \gamma_1\gamma_2}{1 + \mu_1\mu_2} \\
& \leq \frac{1 - (\mu_1 + \mu_2 + \mu_1\mu_2 + \mu_1\gamma_2 + \mu_2\gamma_1)}{1 + \mu_1\mu_2} \leq 1,
\end{aligned} \tag{A.1}$$

that is, to say,  $\alpha_1 \oplus_\varepsilon \alpha_2$  is a PFV. Similarly, we can prove that  $\alpha_1 \otimes_\varepsilon \alpha_2$  is also a PFV.

We now prove that  $\lambda_\cdot \alpha$  is a PFV. To achieve this purpose, we first prove that  $m_\cdot \alpha$  is a PFV for any arbitrary positive integer. Let  $m$  be any positive integer and  $\alpha$  is a PFV, then

$$m_\cdot \alpha = \overbrace{\alpha \oplus_\varepsilon \alpha \oplus_\varepsilon \dots \oplus_\varepsilon \alpha}^m. \tag{A.2}$$

Now, we prove that equation (7) holds for all positive integers with induction method.

First of all, we prove that equation (7) holds for  $m = 1$ . Since

$$\begin{aligned}
1_\cdot \alpha &= \left( \frac{(1 + \mu)^1 - (1 - \mu)^1}{(1 + \mu)^1 + (1 - \mu)^1}, \frac{2\nu^1}{(2 - \nu)^1 + \nu^1}, \frac{2\gamma^1}{(2 - \gamma)^1 + \gamma^1} \right) \\
&= (\mu, \nu, \gamma) = \alpha.
\end{aligned} \tag{A.3}$$

Therefore, equation (7) holds for  $m = 1$ .

Second, if equation (7) holds for  $m = k$ , that is to say

$$k_\cdot \alpha = \overbrace{\alpha \oplus_\varepsilon \alpha \oplus_\varepsilon \dots \oplus_\varepsilon \alpha}^k. \tag{A.4}$$

When  $m = k + 1$ , we have the following:

$$\begin{aligned}
& \overbrace{\alpha \oplus_\varepsilon \alpha \oplus_\varepsilon \dots \oplus_\varepsilon \alpha}^{k+1} = \overbrace{\alpha \oplus_\varepsilon \alpha \oplus_\varepsilon \dots \oplus_\varepsilon \alpha \oplus_\varepsilon \alpha}^{k+1}, \\
& = \left( \frac{(1 + \mu)^k - (1 - \mu)^k}{(1 + \mu)^k + (1 - \mu)^k}, \frac{2\nu^k}{(2 - \nu)^k + \nu^k}, \frac{2\gamma^k}{(2 - \gamma)^k + \gamma^k} \right) \oplus_\varepsilon (\mu, \nu, \gamma), \\
& = \left( \frac{(1 + \mu)^k - (1 - \mu)^k}{(1 + \mu)^k + (1 - \mu)^k}, \frac{2\nu^k}{(2 - \nu)^k + \nu^k}, \frac{2\gamma^k}{(2 - \gamma)^k + \gamma^k} \right) \oplus_\varepsilon \left( \frac{(1 + \mu) - (1 - \mu)}{(1 + \mu) + (1 - \mu)}, \frac{2\nu}{(2 - \nu) + \nu}, \frac{2\gamma}{(2 - \gamma) + \gamma} \right), \\
& \quad \left( \frac{2\gamma^k / (2 - \gamma)^k + \gamma^k \times 2\gamma / (2 - \gamma) + \gamma}{1 + (1 - 2\gamma^k / (2 - \gamma)^k + \gamma^k)(1 - 2\gamma / (2 - \gamma) + \gamma)} \right), \\
& \quad \left( \frac{2\nu^k / (2 - \nu)^k + \nu^k \times 2\nu / (2 - \nu) + \nu}{1 + (1 - 2\nu^k / (2 - \nu)^k + \nu^k)(1 - 2\nu / (2 - \nu) + \nu)} \right), \\
& \quad \left( \frac{2\gamma^k / (2 - \gamma)^k + \gamma^k \times 2\gamma / (2 - \gamma) + \gamma}{1 + (1 - 2\gamma^k / (2 - \gamma)^k + \gamma^k)(1 - 2\gamma / (2 - \gamma) + \gamma)} \right) \\
& = \left( \frac{(1 + \mu)^{k+1} - (1 - \mu)^{k+1}}{(1 + \mu)^{k+1} + (1 - \mu)^{k+1}}, \frac{2\nu^{k+1}}{(2 - \nu)^{k+1} + \nu^{k+1}}, \frac{2\gamma^{k+1}}{(2 - \gamma)^{k+1} + \gamma^{k+1}} \right) \\
& = (k + 1)_\cdot \alpha
\end{aligned} \tag{A.5}$$

That is, to say that equation (7) holds for  $m = k + 1$ . Therefore, equation (7) holds for all positive integers.

Now, we prove equation (7) holds for all positive real numbers. Let  $\lambda$  be any positive real number. Since

$0 \leq \mu, \nu, \gamma \leq 1$  and  $0 \leq \mu + \nu + \gamma \leq 1$ , we obtain  $1 - \mu \geq \nu$ ,  $1 - \mu \geq \gamma$ . Thus, we have  $1 - \mu \geq \min(\nu, \gamma)$ . We also have  $\nu \geq \min(\nu, \gamma)$  and  $\gamma \geq \min(\nu, \gamma)$ . Knowing that

$$\begin{aligned} \frac{(1 + \mu)^\lambda - (1 - \mu)^\lambda}{(1 + \mu)^\lambda + (1 - \mu)^\lambda} &\leq 1 - \frac{2(1 - \mu)^\lambda}{(1 + \mu)^\lambda + (1 - \mu)^\lambda} \leq 1 - \frac{2(1 - \mu)^\lambda}{(1 + \mu)^\lambda + [\min(\nu, \gamma)]^\lambda} \\ \frac{2\nu^\lambda}{(2 - \nu)^\lambda + \nu^\lambda} &\leq \frac{2\nu^\lambda}{(1 + \mu)^\lambda + [\min(\nu, \gamma)]^\lambda}, \end{aligned} \quad (\text{A.6})$$

and

$$\frac{2\gamma^\lambda}{(2 - \gamma)^\lambda + \gamma^\lambda} \leq \frac{2\gamma^\lambda}{(1 + \mu)^\lambda + [\min(\nu, \gamma)]^\lambda}, \quad (\text{A.7})$$

we have

$$\begin{aligned} &\frac{(1 + \mu)^\lambda - (1 - \mu)^\lambda}{(1 + \mu)^\lambda + (1 - \mu)^\lambda} + \frac{2\nu^\lambda}{(2 - \nu)^\lambda + \nu^\lambda} + \frac{2\gamma^\lambda}{(2 - \gamma)^\lambda + \gamma^\lambda} \\ &= \frac{(1 + \mu)^\lambda + (1 - \mu)^\lambda - 2(1 - \mu)^\lambda}{(1 + \mu)^\lambda + (1 - \mu)^\lambda} + \frac{2\nu^\lambda}{(2 - \nu)^\lambda + \nu^\lambda} + \frac{2\gamma^\lambda}{(2 - \gamma)^\lambda + \gamma^\lambda} \\ &= 1 - \frac{2(1 - \mu)^\lambda}{(1 + \mu)^\lambda + (1 - \mu)^\lambda} + \frac{2\nu^\lambda}{(2 - \nu)^\lambda + \nu^\lambda} + \frac{2\gamma^\lambda}{(2 - \gamma)^\lambda + \gamma^\lambda} \\ &\leq 1 - \frac{2(1 - \mu)^\lambda}{(1 + \mu)^\lambda + [\min(\nu, \gamma)]^\lambda} + \frac{2\nu^\lambda}{(1 + \mu)^\lambda + [\min(\nu, \gamma)]^\lambda} + \frac{2\gamma^\lambda}{(1 + \mu)^\lambda + [\min(\nu, \gamma)]^\lambda} \\ &\leq 1 - \frac{2(\nu + \gamma)^\lambda - 2\nu^\lambda - 2\gamma^\lambda}{(1 + \mu)^\lambda + [\min(\nu, \gamma)]^\lambda}. \end{aligned} \quad (\text{A.8})$$

Since

$$\begin{aligned} &2(\nu + \gamma)^\lambda - 2\nu^\lambda - 2\gamma^\lambda \\ &= 2[(\nu^\lambda + \nu^{\lambda-1}\gamma + \nu^{\lambda-2}\gamma^2 + \dots + \nu\gamma^{\lambda-1} + \gamma^\lambda) - \nu^\lambda - \gamma^\lambda] \\ &= 2(\nu^{\lambda-1}\gamma + \nu^{\lambda-2}\gamma^2 + \dots + \nu\gamma^{\lambda-1}) \geq 0, \end{aligned} \quad (\text{A.9})$$

we have

$$\begin{aligned} &1 - \frac{2(\nu + \gamma)^\lambda - 2\nu^\lambda - 2\gamma^\lambda}{(1 + \mu)^\lambda + [\min(\nu, \gamma)]^\lambda} \\ &\leq 1 - \frac{2(\nu^{\lambda-1}\gamma + \nu^{\lambda-2}\gamma^2 + \dots + \nu\gamma^{\lambda-1})}{(1 + \mu)^\lambda + [\min(\nu, \gamma)]^\lambda} \leq 1. \end{aligned} \quad (\text{A.10})$$

Furthermore,

$$\begin{aligned} \frac{(1+\mu)^\lambda - (1-\mu)^\lambda}{(1+\mu)^\lambda + (1-\mu)^\lambda} &\geq 0, \\ \frac{2\nu^\lambda}{(2-\nu)^\lambda + \nu^\lambda} &\geq 0, \\ \frac{2\gamma^\lambda}{(2-\gamma)^\lambda + \gamma^\lambda} &\geq 0. \end{aligned} \quad (\text{A.11})$$

With the above analysis, we have

$$0 \leq \frac{(1+\mu)^\lambda - (1-\mu)^\lambda}{(1+\mu)^\lambda + (1-\mu)^\lambda} + \frac{2\nu^\lambda}{(2-\nu)^\lambda + \nu^\lambda} + \frac{2\gamma^\lambda}{(2-\gamma)^\lambda + \gamma^\lambda} \leq 1. \quad (\text{A.12})$$

That is, to say,  $\alpha_1 \otimes_\varepsilon \alpha_2$  is also a PFV for any positive real number.

Similar to  $\lambda_\varepsilon \alpha$ , we can prove that  $\alpha^\lambda$  is a PFV.  $\square$

## B. Proof of Theorem 4

*Proof:* First, according to Theorem 3, it is clear that the aggregated value with picture fuzzy Einstein hybrid-weighted average (PFEHWA) operator is also a PFV.

We now prove that equation (17) holds. According to the operations of PFVs defined in Definition 8, we have

$$\begin{aligned} \frac{\omega_{\varepsilon(j)}\lambda_j}{\sum_{j=1}^n \omega_{\varepsilon(j)}\lambda_j} \alpha_j &= \left( \frac{\left( (1+\mu_{\alpha_j})^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)}\lambda_j} - (1-\mu_{\alpha_j})^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)}\lambda_j} \right)}{\left( (1+\mu_{\alpha_j})^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)}\lambda_j} + (1-\mu_{\alpha_j})^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)}\lambda_j} \right)} \right. \\ &\quad \left. \frac{2\nu_{\alpha_j}^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)}\lambda_j}}{\left( 2-\nu_{\alpha_j} \right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)}\lambda_j} + \nu_{\alpha_j}^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)}\lambda_j}}, \frac{2\gamma_{\alpha_j}^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)}\lambda_j}}{\left( 2-\gamma_{\alpha_j} \right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)}\lambda_j} + \gamma_{\alpha_j}^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)}\lambda_j}} \right), \end{aligned} \quad (\text{B.1})$$

where  $j = 1, 2, \dots, n$ . We prove the correctness of equation (17) with inductive method.

(1) for  $n = 1$ , it is trivial.

(2) Suppose equation (17) holds for  $n = k$ , that is:

$$\begin{aligned} \text{PFEHWA}_{\omega, \lambda}(\alpha_1, \alpha_2, \dots, \alpha_k) &= \left( \frac{\prod_{j=1}^k (1+\mu_{\alpha_j})^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^k \omega_{\varepsilon(j)}\lambda_j} - \prod_{j=1}^k (1-\mu_{\alpha_j})^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^k \omega_{\varepsilon(j)}\lambda_j}}{\prod_{j=1}^k (1+\mu_{\alpha_j})^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^k \omega_{\varepsilon(j)}\lambda_j} + \prod_{j=1}^k (1-\mu_{\alpha_j})^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^k \omega_{\varepsilon(j)}\lambda_j}} \right. \\ &\quad \frac{2 \prod_{j=1}^k \nu_{\alpha_j}^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^k \omega_{\varepsilon(j)}\lambda_j}}{\prod_{j=1}^k \left( 2-\nu_{\alpha_j} \right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^k \omega_{\varepsilon(j)}\lambda_j} + \prod_{j=1}^k \nu_{\alpha_j}^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^k \omega_{\varepsilon(j)}\lambda_j}} \\ &\quad \left. \frac{2 \prod_{j=1}^k \gamma_{\alpha_j}^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^k \omega_{\varepsilon(j)}\lambda_j}}{\prod_{j=1}^k \left( 2-\gamma_{\alpha_j} \right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^k \omega_{\varepsilon(j)}\lambda_j} + \prod_{j=1}^k \gamma_{\alpha_j}^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^k \omega_{\varepsilon(j)}\lambda_j}} \right). \end{aligned} \quad (\text{B.2})$$

(3) When  $n = k + 1$ , according to Definition 9 and Theorem 3, we have as follows:



$$= \left( \frac{(1 + \mu_A) \sum_{j=1}^m \omega_{\varepsilon(j)\lambda_j} / \sum_{j=1}^n \omega_{\varepsilon(j)\lambda_j} - (1 - \mu_A) \sum_{j=1}^m \omega_{\varepsilon(j)\lambda_j} / \sum_{j=1}^n \omega_{\varepsilon(j)\lambda_j}}{(1 + \mu_{A_j}) \sum_{j=1}^m \omega_{\varepsilon(j)\lambda_j} / \sum_{j=1}^n \omega_{\varepsilon(j)\lambda_j} + (1 - \mu_{A_j}) \sum_{j=1}^m \omega_{\varepsilon(j)\lambda_j} / \sum_{j=1}^n \omega_{\varepsilon(j)\lambda_j}}, \frac{2\gamma_A \sum_{j=1}^m \omega_{\varepsilon(j)\lambda_j} / \sum_{j=1}^n \omega_{\varepsilon(j)\lambda_j}}{(2 - \gamma_A) \sum_{j=1}^m \omega_{\varepsilon(j)\lambda_j} / \sum_{j=1}^n \omega_{\varepsilon(j)\lambda_j} + \gamma_A \sum_{j=1}^m \omega_{\varepsilon(j)\lambda_j} / \sum_{j=1}^n \omega_{\varepsilon(j)\lambda_j}} \right) \quad (C.1)$$

## (2) Boundedness

Let  $f(x) = (1-x)/(1+x)$ ,  $x \in [0, 1]$ . Then,  $f'(x) = -2/(1+x)^2 > 0$  and thus,  $f(x)$  is a decreasing function. Since  $\min \mu_{A_j} \leq \mu_{A_j} \leq \max \mu_{A_j}$ , for all  $j$ , then  $f(\max \mu_{A_j}) \leq f(\mu_{A_j}) \leq f(\min \mu_{A_j})$ . Let  $\alpha = \dot{\omega}_{\varepsilon(j)\lambda_j} / \sum_{j=1}^n \omega_{\varepsilon(j)\lambda_j}$ . Therefore, we have as follows:

$$\frac{(1 - \max \mu_{A_j})}{(1 + \max \mu_{A_j})} \leq \frac{(1 - \mu_{A_j})}{(1 + \mu_{A_j})} \leq \frac{(1 - \min \mu_{A_j})}{(1 + \min \mu_{A_j})} \quad (C.2)$$

So, we can derive the following:

$$\frac{(1 - \max \mu_{A_j})^{\omega_{\varepsilon(j)\lambda_j} / \sum_{j=1}^n \omega_{\varepsilon(j)\lambda_j}}}{(1 + \max \mu_{A_j})^{\omega_{\varepsilon(j)\lambda_j} / \sum_{j=1}^n \omega_{\varepsilon(j)\lambda_j}}} \leq \frac{(1 - \mu_{A_j})^{\omega_{\varepsilon(j)\lambda_j} / \sum_{j=1}^n \omega_{\varepsilon(j)\lambda_j}}}{(1 + \mu_{A_j})^{\omega_{\varepsilon(j)\lambda_j} / \sum_{j=1}^n \omega_{\varepsilon(j)\lambda_j}}} \leq \frac{(1 - \min \mu_{A_j})^{\omega_{\varepsilon(j)\lambda_j} / \sum_{j=1}^n \omega_{\varepsilon(j)\lambda_j}}}{(1 + \min \mu_{A_j})^{\omega_{\varepsilon(j)\lambda_j} / \sum_{j=1}^n \omega_{\varepsilon(j)\lambda_j}}} \quad (C.3)$$

Thus,

$$\begin{aligned} \prod_{j=1}^m \frac{(1 - \max \mu_{A_j})^{\omega_{\varepsilon(j)\lambda_j} / \sum_{j=1}^n \omega_{\varepsilon(j)\lambda_j}}}{(1 + \max \mu_{A_j})^{\omega_{\varepsilon(j)\lambda_j} / \sum_{j=1}^n \omega_{\varepsilon(j)\lambda_j}}} &\leq \prod_{j=1}^m \frac{(1 - \mu_{A_j})^{\omega_{\varepsilon(j)\lambda_j} / \sum_{j=1}^m \omega_{\varepsilon(j)\lambda_j}}}{(1 + \mu_{A_j})^{\omega_{\varepsilon(j)\lambda_j} / \sum_{j=1}^m \omega_{\varepsilon(j)\lambda_j}}} \\ &\leq \prod_{j=1}^m \frac{(1 - \min \mu_{A_j})^{\omega_{\varepsilon(j)\lambda_j} / \sum_{j=1}^m \omega_{\varepsilon(j)\lambda_j}}}{(1 + \min \mu_{A_j})^{\omega_{\varepsilon(j)\lambda_j} / \sum_{j=1}^m \omega_{\varepsilon(j)\lambda_j}}} \\ &\Leftrightarrow \frac{(1 - \max \mu_{A_j})^{\sum_{j=1}^m (\omega_{\varepsilon(j)\lambda_j} / \sum_{j=1}^m \omega_{\varepsilon(j)\lambda_j})}}{(1 + \max \mu_{A_j})^{\sum_{j=1}^m (\omega_{\varepsilon(j)\lambda_j} / \sum_{j=1}^m \omega_{\varepsilon(j)\lambda_j})}} \leq \prod_{j=1}^m \frac{(1 - \mu_{A_j})^{\omega_{\varepsilon(j)\lambda_j} / \sum_{j=1}^m \omega_{\varepsilon(j)\lambda_j}}}{(1 + \mu_{A_j})^{\omega_{\varepsilon(j)\lambda_j} / \sum_{j=1}^m \omega_{\varepsilon(j)\lambda_j}}} \\ &\leq \frac{(1 - \min \mu_{A_j})^{\sum_{j=1}^m (\omega_{\varepsilon(j)\lambda_j} / \sum_{j=1}^m \omega_{\varepsilon(j)\lambda_j})}}{(1 + \min \mu_{A_j})^{\sum_{j=1}^m (\omega_{\varepsilon(j)\lambda_j} / \sum_{j=1}^m \omega_{\varepsilon(j)\lambda_j})}} \end{aligned}$$

$$\begin{aligned}
& \frac{1 - \max_j \mu_{A_j}}{1 + \max_j \mu_{A_j}} \leq \prod_{j=1}^m \frac{(1 - \mu_{A_j})^{\omega_{\varepsilon(j)} \lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)} \lambda_j}}{(1 + \mu_{A_j})^{\omega_{\varepsilon(j)} \lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)} \lambda_j}} \leq \frac{1 - \min_j \mu_{A_j}}{1 + \min_j \mu_{A_j}} \\
& \Leftrightarrow \frac{2}{1 + \max_j \mu_{A_j}} \leq 1 + \prod_{j=1}^m \frac{(1 - \mu_{A_j})^{\omega_{\varepsilon(j)} \lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)} \lambda_j}}{(1 + \mu_{A_j})^{\omega_{\varepsilon(j)} \lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)} \lambda_j}} \leq \frac{2}{1 + \min_j \mu_{A_j}} \\
& \Leftrightarrow \frac{1 + \min_j \mu_{A_j}}{2} \leq \frac{1}{1 + \prod_{j=1}^m (1 - \mu_{A_j})^{\omega_{\varepsilon(j)} \lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)} \lambda_j} / (1 + \mu_{A_j})^{\omega_{\varepsilon(j)} \lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)} \lambda_j}} \leq \frac{1 + \max_j \mu_{A_j}}{2} \tag{C.4} \\
& \Leftrightarrow \min_j \mu_{A_j} \leq \frac{2}{1 + \prod_{j=1}^m (1 - \mu_{A_j})^{\omega_{\varepsilon(j)} \lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)} \lambda_j} / (1 + \mu_{A_j})^{\omega_{\varepsilon(j)} \lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)} \lambda_j}} - 1 \leq \max_j \mu_{A_j} \\
& = \min_j \mu_{A_j} \leq \frac{2}{1 + \prod_{j=1}^m (1 - \mu_{A_j})^{\omega_{\varepsilon(j)} \lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)} \lambda_j} / (1 + \mu_{A_j})^{\omega_{\varepsilon(j)} \lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)} \lambda_j}} - 1 \leq \max_j \mu_{A_j} \\
& \Leftrightarrow \min_j \mu_{A_j} \leq \frac{(1 + \mu_{A_j})^{\omega_{\varepsilon(j)} \lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)} \lambda_j} - (1 - \mu_{A_j})^{\omega_{\varepsilon(j)} \lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)} \lambda_j}}{(1 + \mu_{A_j})^{\omega_{\varepsilon(j)} \lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)} \lambda_j} + (1 - \mu_{A_j})^{\omega_{\varepsilon(j)} \lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)} \lambda_j}} \leq \max_j \mu_{A_j}.
\end{aligned}$$

Let  $g(y) = (2 - y)/y$ ,  $y \in [0, 1]$ . Then,  $f'(y) = -2/y^2 < 0$ , and thus,  $g(y)$  is a decreasing function. Since  $\min_j \nu_{A_j} \leq \nu_{A_j} \leq \max_j \nu_{A_j}$  for all  $j$ , then  $g(\min_j \nu_{A_j}) \leq$

$g(\nu_{A_j}) \leq g(\max_j \nu_{A_j})$ . Let  $\alpha = \omega_{\varepsilon(j)} \lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)} \lambda_j$ . According to equation (19), we have as follows:

$$\frac{(2 - \max_j \nu_{A_j})^{\omega_{\varepsilon(j)} \lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)} \lambda_j}}{\max_j \nu_{A_j}^{\omega_{\varepsilon(j)} \lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)} \lambda_j}} \leq \frac{(2 - \nu_{A_j})^{\omega_{\varepsilon(j)} \lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)} \lambda_j}}{\nu_{A_j}^{\omega_{\varepsilon(j)} \lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)} \lambda_j}} \leq \frac{(2 - \min_j \nu_{A_j})^{\omega_{\varepsilon(j)} \lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)} \lambda_j}}{\min_j \nu_{A_j}^{\omega_{\varepsilon(j)} \lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)} \lambda_j}} \tag{C.5}$$

So, we can derive the following:

$$\begin{aligned}
& \frac{\prod_{j=1}^m \left(2 - \max_j v_{A_j}\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)}\lambda_j}}{\prod_{j=1}^m \max_j v_{A_j}^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)}\lambda_j}} \leq \frac{\prod_{j=1}^m \left(2 - v_{A_j}\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)}\lambda_j}}{\prod_{j=1}^m v_{A_j}^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)}\lambda_j}} \leq \frac{\prod_{j=1}^m \left(2 - \min_j v_{A_j}\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)}\lambda_j}}{\prod_{j=1}^m \min_j v_{A_j}^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)}\lambda_j}} \\
& \Leftrightarrow \frac{2 - \max_j v_{A_j}}{\max_j v_{A_j}} \leq \frac{\prod_{j=1}^m \left(2 - v_{A_j}\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)}\lambda_j}}{\prod_{j=1}^m v_{A_j}^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)}\lambda_j}} \leq \frac{2 - \min_j v_{A_j}}{\min_j v_{A_j}} \tag{C.6} \\
& \Leftrightarrow \frac{2}{\max_j v_{A_j}} \leq \frac{\prod_{j=1}^m \left(2 - v_{A_j}\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)}\lambda_j}}{\prod_{j=1}^m v_{A_j}^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)}\lambda_j}} + 1 \leq \frac{2}{\min_j v_{A_j}} \\
& \Leftrightarrow \min_j v_{A_j} \leq \frac{2 \prod_{j=1}^m v_{A_j}^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)}\lambda_j}}{\prod_{j=1}^m \left(2 - v_{A_j}\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)}\lambda_j} + \prod_{j=1}^m v_{A_j}^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)}\lambda_j}} \leq \max_j v_{A_j}.
\end{aligned}$$

Similarly, we have as follows:

$$\min_j \gamma_{A_j} \leq \frac{2 \prod_{j=1}^m \gamma_{A_j}^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)}\lambda_j}}{\prod_{j=1}^m \left(2 - \gamma_{A_j}\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)}\lambda_j} + \prod_{j=1}^m \gamma_{A_j}^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^n \omega_{\varepsilon(j)}\lambda_j}} \leq \max_j \gamma_{A_j}. \tag{C.7}$$

According to Definition 4, we can have  $A_{\min} \subseteq A \subseteq A_{\max}$ .

### (3) Monotonicity

Let  $A_j^1 = (\mu_{A_j}^1, \nu_{A_j}^1, \gamma_{A_j}^1)$  and  $A_j^2 = (\mu_{A_j}^2, \nu_{A_j}^2, \gamma_{A_j}^2)$  be two collections of PFVs, and  $\mu_{A_j}^1 \leq \mu_{A_j}^2$ ,  $\nu_{A_j}^1 \geq \nu_{A_j}^2$ , and  $\gamma_{A_j}^1 \geq \gamma_{A_j}^2$ . Let  $\text{PFEHWA}_{\omega, \lambda}(A_1^1, A_2^1, \dots, A_m^1) = A^1 = (\mu_{A^1}, \nu_{A^1}, \gamma_{A^1})$  and  $\text{PFEHWA}_{\omega, \lambda}(A_1^2, A_2^2, \dots, A_m^2) = A^2 = (\mu_{A^2}, \nu_{A^2}, \gamma_{A^2})$ . Let  $f(x) = (1+x)/(1-x)$ ,  $x \in [0, 1]$ ,  $\alpha \in [0, 1]$ . Thus,  $f(x)$  is an increasing function. If  $\mu_{A_j}^1 \leq \mu_{A_j}^2$  for all  $j$ , then  $f(\mu_{A_j}^1) \leq f(\mu_{A_j}^2)$ , i.e.,  $(1 + \mu_{A_j}^1)/(1 - \mu_{A_j}^1) \leq (1 + \mu_{A_j}^2)/(1 - \mu_{A_j}^2)$ . Let  $\alpha = \omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)}\lambda_j$ . Therefore, we have as follows:

$$\begin{aligned}
& \frac{\prod_{i=1}^m \left(1 + \mu_{A_j}^1\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)}\lambda_j}}{\prod_{i=1}^m \left(1 - \mu_{A_j}^1\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)}\lambda_j}} \leq \frac{\prod_{i=1}^m \left(1 + \mu_{A_j}^2\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)}\lambda_j}}{\prod_{i=1}^m \left(1 - \mu_{A_j}^2\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)}\lambda_j}} \\
& \Leftrightarrow 1 + \frac{\prod_{i=1}^m \left(1 + \mu_{A_j}^1\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)}\lambda_j}}{\prod_{i=1}^m \left(1 - \mu_{A_j}^1\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)}\lambda_j}} \leq 1 + \frac{\prod_{i=1}^m \left(1 + \mu_{A_j}^2\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)}\lambda_j}}{\prod_{i=1}^m \left(1 - \mu_{A_j}^2\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)}\lambda_j}} \tag{C.8}
\end{aligned}$$

So, we can derive the following:

$$\begin{aligned}
& \Leftrightarrow \frac{2}{1 + \prod_{i=1}^m \left(1 + \mu_{A_j}^2\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)}\lambda_j} / \prod_{i=1}^m \left(1 - \mu_{A_j}^2\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)}\lambda_j}} \\
& \leq \frac{2}{1 + \prod_{i=1}^m \left(1 + \mu_{A_j}^1\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)}\lambda_j} / \prod_{i=1}^m \left(1 - \mu_{A_j}^1\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)}\lambda_j}} \\
& \Leftrightarrow 1 - \frac{2}{1 + \prod_{i=1}^m \left(1 + \mu_{A_j}^1\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)}\lambda_j} / \prod_{i=1}^m \left(1 - \mu_{A_j}^1\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)}\lambda_j}} \\
& \leq 1 - \frac{2}{1 + \prod_{i=1}^m \left(1 + \mu_{A_j}^2\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)}\lambda_j} / \prod_{i=1}^m \left(1 - \mu_{A_j}^2\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)}\lambda_j}} \tag{C.9} \\
& \Leftrightarrow \frac{\prod_{i=1}^m \left(1 + \mu_{A_j}^1\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)}\lambda_j} - \prod_{i=1}^m \left(1 - \mu_{A_j}^1\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)}\lambda_j}}{\prod_{i=1}^m \left(1 + \mu_{A_j}^1\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)}\lambda_j} + \prod_{i=1}^m \left(1 - \mu_{A_j}^1\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)}\lambda_j}} \\
& \leq \frac{\prod_{i=1}^m \left(1 + \mu_{A_j}^2\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)}\lambda_j} - \prod_{i=1}^m \left(1 - \mu_{A_j}^2\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)}\lambda_j}}{\prod_{i=1}^m \left(1 + \mu_{A_j}^2\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)}\lambda_j} + \prod_{i=1}^m \left(1 - \mu_{A_j}^2\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)}\lambda_j}} \\
& \Leftrightarrow \mu_{A^1} \leq \mu_{A^2}.
\end{aligned}$$

Let  $g(y) = (2 - y)/y$ ,  $y \in [0, 1]$ ,  $\alpha \in [0, 1]$ . Moreover, thus  $g(y)$  is an increasing function. Since  $\nu_{A_j^1} \geq \nu_{A_j^2}$  for all  $j$ , then  $g(\nu_{A_j^1}) \leq g(\nu_{A_j^2})$ , i.e.,  $(2 - \nu_{A_j^1})/\nu_{A_j^1} \leq (2 - \nu_{A_j^2})/\nu_{A_j^2}$ . Let

$\alpha = \omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)}\lambda_j$ . According to equation (20), we have as follows:

$$\begin{aligned}
& \left(2 - \nu_{A_j^1}/\nu_{A_j^1}\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)}\lambda_j} \leq \left(2 - \nu_{A_j^2}/\nu_{A_j^2}\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)}\lambda_j} \\
& \Leftrightarrow \prod_{j=1}^m \left(2 - \nu_{A_j^1}/\nu_{A_j^1}\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)}\lambda_j} \leq \prod_{j=1}^m \left(2 - \nu_{A_j^2}/\nu_{A_j^2}\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)}\lambda_j} \\
& \Leftrightarrow 1 + \prod_{j=1}^m \left(2 - \nu_{A_j^1}/\nu_{A_j^1}\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)}\lambda_j} \leq 1 + \prod_{j=1}^m \left(2 - \nu_{A_j^2}/\nu_{A_j^2}\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)}\lambda_j} \\
& \Leftrightarrow \frac{2}{1 + \prod_{j=1}^m \left(2 - \nu_{A_j^2}/\nu_{A_j^2}\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)}\lambda_j}} \leq \frac{2}{1 + \prod_{j=1}^m \left(2 - \nu_{A_j^1}/\nu_{A_j^1}\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)}\lambda_j}} \\
& \Leftrightarrow \frac{\prod_{j=1}^m \left(\nu_{A_j^2}\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)}\lambda_j}}{\prod_{j=1}^m \left(2 - \nu_{A_j^2}\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)}\lambda_j} + \prod_{j=1}^m \left(\nu_{A_j^2}\right)^{\omega_{\varepsilon(j)}\lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)}\lambda_j}}
\end{aligned}$$



$$\begin{aligned} & \frac{\prod_{j=1}^m \left( \nu_{A_j^1} \right)^{\omega_{\varepsilon(j)} \lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)} \lambda_j}}{\prod_{j=1}^m \left( 2 - \nu_{A_j^1} \right)^{\omega_{\varepsilon(j)} \lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)} \lambda_j} + \prod_{j=1}^m \left( \nu_{A_j^1} \right)^{\omega_{\varepsilon(j)} \lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)} \lambda_j}} \\ & \Leftrightarrow \nu_{A_j^2} \leq \nu_{A_j^1}. \end{aligned} \quad (\text{C.10})$$

Similarly, we have the following:

$$\begin{aligned} & \frac{\prod_{j=1}^m \left( \gamma_{A_j^2} \right)^{\omega_{\varepsilon(j)} \lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)} \lambda_j}}{\prod_{j=1}^m \left( 2 - \gamma_{A_j^2} \right)^{\omega_{\varepsilon(j)} \lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)} \lambda_j} + \prod_{j=1}^m \left( \gamma_{A_j^2} \right)^{\omega_{\varepsilon(j)} \lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)} \lambda_j}} \\ & \leq \frac{\prod_{j=1}^m \left( \gamma_{A_j^1} \right)^{\omega_{\varepsilon(j)} \lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)} \lambda_j}}{\prod_{j=1}^m \left( 2 - \gamma_{A_j^1} \right)^{\omega_{\varepsilon(j)} \lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)} \lambda_j} + \prod_{j=1}^m \left( \gamma_{A_j^1} \right)^{\omega_{\varepsilon(j)} \lambda_j / \sum_{j=1}^m \omega_{\varepsilon(j)} \lambda_j}} \\ & \Leftrightarrow \gamma_{A_j^2} \leq \gamma_{A_j^1}. \end{aligned} \quad (\text{C.11})$$

According to Definition 4, we have  $A^1 \subseteq A^2$ , which completes the proof of Theorem 5.  $\square$

## Data Availability

The data used to support the findings of this study are included within the article.

## Conflicts of Interest

The author declares that there are no conflicts of interest regarding the publication of this paper.

## Acknowledgments

This research was partially supported by The Jiaoyuting Gaoxiao Social Science Foundation of Jiangsu Province (No. 2017ZDIXM009) during which the study was completed.

## References

- [1] M. Tavana and D. Di Caprio, "Modeling synergies in multicriteria supplier selection and order allocation: an application to commodity trading," *European Journal of Operational Research*, vol. 254, pp. 859–874, 2016.
- [2] A. Chauhan and A. Singh, "A hybrid multi-criteria decision making method approach for selecting a sustainable location of healthcare waste disposal facility," *Journal of Cleaner Production*, vol. 139, no. 15, pp. 1001–1010, 2016.
- [3] H. Liao, Z. Xu, X.-J. Zeng, and D.-L. Xu, "An enhanced consensus reaching process in group decision making with intuitionistic fuzzy preference relations," *Information Sciences*, vol. 329, pp. 274–286, 2016.
- [4] P. Gupta, M. K. Mehlatat, and N. Grover, "Intuitionistic fuzzy multi-attribute group decision-making with an application to plant location selection based on a new extended VIKOR method," *Information Sciences*, vol. 370–371, pp. 184–203, 2016.
- [5] F. Meng, S.-M. Chen, and J. Tang, "Multicriteria decision making based on bi-direction Choquet integrals," *Information Sciences*, vol. 555, pp. 339–356, 2021.
- [6] M. Juanpera, B. Domenech, L. Ferrer-Martí, A. García-Villoria, and R. Pastor, "Methodology for integrated multicriteria decision-making with uncertainty: extending the compromise ranking method for uncertain evaluation of alternatives," *Fuzzy Sets and Systems*, vol. 434, pp. 135–158, 2022.
- [7] S. K. Ghosh, A. Mitra, and Ghosh, "A novel intuitionistic fuzzy soft set entrenched mammogram segmentation under Multi-granulation approximation for breast cancer detection in early stages," *Expert Systems with Applications*, vol. 169, Article ID 114329, 2020.
- [8] Y. Jiang, Z. Xu, and X. Yu, "Group decision making based on incomplete intuitionistic multiplicative preference relations," *Information Sciences*, vol. 295, pp. 33–52, 2015.
- [9] F. Chiclana, F. Herrera, and E. Herrera-Viedma, "Integrating three representation models in fuzzy multipurpose decision making based on fuzzy preference relations," *Fuzzy Sets and Systems*, vol. 97, no. 1, pp. 33–48, 1998.
- [10] Y.-M. Wang and C. Parkan, "Multiple attribute decision making based on fuzzy preference information on alternatives: ranking and weighting," *Fuzzy Sets and Systems*, vol. 153, no. 3, pp. 331–346, 2005.
- [11] F. E. Boran and D. Akay, "A biparametric similarity measure on intuitionistic fuzzy sets with applications to pattern recognition," *Information Sciences*, vol. 255, pp. 45–57, 2014.
- [12] F. Shen, J. Xu, and Z. Xu, "An outranking sorting method for multi-criteria group decision making using intuitionistic fuzzy sets," *Information Sciences*, vol. 334–335, pp. 338–353, 2016.
- [13] S.-M. Chen and C.-H. Chang, "Fuzzy multiattribute decision making based on transformation techniques of intuitionistic

- fuzzy values and intuitionistic fuzzy geometric averaging operators,” *Information Sciences*, vol. 352-353, no. 353, pp. 133-149, 2016.
- [14] S.-M. Chen, S.-H. Cheng, and T.-C. Lan, “Multicriteria decision making based on the TOPSIS method and similarity measures between intuitionistic fuzzy values,” *Information Sciences*, vol. 367-368, pp. 279-295, 2016.
- [15] Z. S. Xu and R. R. Yager, “Some geometric aggregation operators based on intuitionistic fuzzy sets,” *International Journal of General Systems*, vol. 35, no. 6, pp. 417-433, 2006.
- [16] V. Başhan, H. Demirel, and M. Gul, “An FMEA-based TOPSIS approach under single valued neutrosophic sets for maritime risk evaluation: the case of ship navigation safety,” *Soft Computing*, vol. 24, pp. 18749-18764, 2020.
- [17] R.-p. Tan and W.-d. Zhang, “Decision-making method based on new entropy and refined single-valued neutrosophic sets and its application in typhoon disaster assessment,” *Applied Intelligence*, vol. 51, no. 1, pp. 283-307, 2021.
- [18] E. Bolturk and C. Kahraman, “A novel interval-valued neutrosophic AHP with cosine similarity measure,” *Soft Computing*, vol. 22, no. 15, pp. 4941-4958, 2018.
- [19] J. Mahanta and S. Panda, “Distance measure for Pythagorean fuzzy sets with varied applications,” *Neural Computing & Applications*, vol. 33, no. 24, pp. 17161-17171, 2021.
- [20] M. Akram, F. Ilyas, and H. Garg, “Multi-criteria group decision making based on ELECTRE I method in Pythagorean fuzzy information,” *Soft Computing*, vol. 24, no. 5, pp. 3425-3453, 2020.
- [21] T. Senapati, R. R. Yager, “Fermatean fuzzy sets,” *Journal of Ambient Intelligence and Humanized Computing*, vol. 11, no. 2, pp. 663-674, 2020.
- [22] C. Shit and G. Ghorai, “Multiple attribute decision-making based on different types of Dombi aggregation operators under Fermatean fuzzy information,” *Soft Computing*, vol. 25, no. 22, pp. 13869-13880, 2021.
- [23] F. Zhang, S. Chen, J. Li, and W. Huang, “New distance measures on hesitant fuzzy sets based on the cardinality theory and their application in pattern recognition,” *Soft Computing*, vol. 22, no. 4, pp. 1237-1245, 2018.
- [24] Z. Hussain and M.-S. Yang, “Entropy for hesitant fuzzy sets based on hausdorff metric with construction of hesitant fuzzy TOPSIS,” *International Journal of Fuzzy Systems*, vol. 20, no. 8, pp. 2517-2533, 2018.
- [25] W. Zeng, D. Li, and Q. Yin, “Weighted interval-valued hesitant fuzzy sets and its application in group decision making,” *International Journal of Fuzzy Systems*, vol. 21, no. 2, pp. 421-432, 2019.
- [26] K. T. Atanassov, “Intuitionistic fuzzy sets,” *Fuzzy Sets and Systems*, vol. 20, no. 1, pp. 87-96, 1986.
- [27] F. Smarandache, *A Unifying Field in Logics: Neutrosophy: Neutrosophic Probability Set and Logics*, American Research Press, Rehoboth, 1999.
- [28] H. W. F. S. Y. Zhang and R. Sunderraman, “Single valued neutrosophic sets,” *Multispace Multistruct*, vol. 4, pp. 410-413, 2012.
- [29] B. C. Cuong and V. Kreinovich, “Picture fuzzy sets,” *Journal of Computer Science & Cybernetics*, vol. 30, pp. 409-416, 2014.
- [30] L. H. Son, “Generalized picture distance measure and applications to picture fuzzy clustering,” *Applied Soft Computing*, vol. 46, pp. 284-295, 2016.
- [31] P. Singh, “Correlation coefficients for picture fuzzy sets,” *Journal of Intelligent and Fuzzy Systems*, vol. 28, no. 2, pp. 591-604, 2015.
- [32] P. H. Thong and L. H. Son, “A novel automatic picture fuzzy clustering method based on particle swarm optimization and picture composite cardinality,” *Knowledge-Based Systems*, vol. 109, pp. 48-60, 2016.
- [33] H. Liao and Z. Xu, “Intuitionistic fuzzy hybrid weighted aggregation operators,” *International Journal of Intelligent Systems*, vol. 29, no. 11, pp. 971-993, 2014.
- [34] Z. Xu and N. Zhao, “Information fusion for intuitionistic fuzzy decision making: an overview,” *Information Fusion*, vol. 28, pp. 10-23, 2016.
- [35] M. Xia and Z. Xu, “Hesitant fuzzy information aggregation in decision making,” *International Journal of Approximate Reasoning*, vol. 52, no. 3, pp. 395-407, 2011.
- [36] S.-M. Chen and W.-H. Tsai, “Multiple attribute decision making based on novel interval-valued intuitionistic fuzzy geometric averaging operators,” *Information Sciences*, vol. 367-368, pp. 1045-1065, 2016.
- [37] C. Jana and M. Pal, “Multi-criteria decision making process based on some single-valued neutrosophic Dombi power aggregation operators,” *Soft Computing*, vol. 25, pp. 5055-5072, 2021.
- [38] H. M. A. Farid and M. Riaz, “Single-valued neutrosophic Einstein interactive aggregation operators with applications for material selection in engineering design: case study of cryogenic storage tank,” *Complex & Intelligent Systems*, DOI: 10.1007/s40747-021-00626-0, 2022.
- [39] H. Garg, “Some picture fuzzy aggregation operators and their applications to multicriteria decision-making,” *Arabian Journal for Science and Engineering*, vol. 42, no. 12, pp. 5275-5290, 2017.
- [40] C. Jana, T. Senapati, M. Pal, and R. R. Yager, “Picture fuzzy Dombi aggregation operators: application to MADM process,” *Applied Soft Computing*, vol. 74, pp. 99-109, 2019.
- [41] S. Khan, S. Abdullah, and S. Ashraf, “Picture fuzzy aggregation information based on Einstein operations and their application in decision making,” *Mathematical Sciences*, vol. 13, no. 3, pp. 213-229, 2019.
- [42] M. Xia, Z. Xu, and B. Zhu, “Some issues on intuitionistic fuzzy aggregation operators based on Archimedean t-conorm and t-norm,” *Knowledge-Based Systems*, vol. 31, pp. 78-88, 2012.
- [43] X. Jia and Y. M. Wang, “Choquet integral-based intuitionistic fuzzy arithmetic aggregation operators in multi-criteria decision-making,” *Expert Systems with Applications*, vol. 191, Article ID 116242, 2022.
- [44] Z. Zhang, “Multi-criteria group decision-making methods based on new intuitionistic fuzzy Einstein hybrid weighted aggregation operators,” *Neural Computing & Applications*, vol. 28, no. 12, pp. 3781-3800, 2017.
- [45] J. Ye, “A multicriteria decision-making method using aggregation operators for simplified neutrosophic sets,” *Journal of Intelligent and Fuzzy Systems*, vol. 26, no. 5, pp. 2459-2466, 2014.
- [46] W. Z. Wang and X. W. Liu, “Some operations over Atanassov’s intuitionistic fuzzy sets based on Einstein T-Norm and T-Conorm,” *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 21, no. 2, pp. 263-276, 2013.
- [47] W. Wang and X. Liu, “Intuitionistic fuzzy geometric aggregation operators based on Einstein operations,” *International Journal of Intelligent Systems*, vol. 26, no. 11, pp. 1049-1075, 2011.
- [48] P. Liu, “The aggregation operators based on Archimedean T-conorm and T-norm for single-valued Neutrosophic numbers

## Retraction

# Retracted: Random Forest Algorithm Based on Speech for Early Identification of Parkinson's Disease

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] P. Fan, "Random Forest Algorithm Based on Speech for Early Identification of Parkinson's Disease," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3287068, 6 pages, 2022.

## Research Article

# Random Forest Algorithm Based on Speech for Early Identification of Parkinson's Disease

**Ping Fan** 

*School of Chinese Language and Literature, Nanjing Normal University, Nanjing, China*

Correspondence should be addressed to Ping Fan; [fpshida2010@126.com](mailto:fpshida2010@126.com)

Received 9 January 2022; Accepted 4 February 2022; Published 9 May 2022

Academic Editor: Daqing Gong

Copyright © 2022 Ping Fan. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

To investigate the effectiveness of identifying patients with Parkinson's disease (PD) from speech signals, various acoustic parameters including prosodic and segmental features are extracted from speech and then the random forest classification (RF) algorithm based on these acoustic parameters is applied to diagnose early-stage PD patients. To validate the proposed method of RF algorithm in early-stage PD identification, this study compares the accuracy rate of RF with that of neurologists' judgments based on auditory test outcomes, and the results clearly show the superiority of the proposed method over its rival. Random forest algorithm based on speech can improve the accuracy of patients' identification, which provides an efficient auxiliary method in the early diagnosis of PD patients.

## 1. Introduction

Parkinson's disease (PD) is a progressive neurodegenerative disease with ambiguous etiology. Although the exact reason is still unknown, the riskiest factor of PD is age, which causes the prevalence rate of PD to increase with the increase of age. The main pathological changes in patients with PD are the death and loss of dopaminergic neurons in the substantia nigra pars compacta which is irreversible. It means that patients with PD will never be cured, but anti-Parkinson's medication or deep brain stimulation surgery can slow down the progression of the disease [1]. So, early detection, early intervention, and early treatment of PD patients are essential to alleviate their pain and the burden of their families.

Patients with PD are usually diagnosed based on clinical symptoms (e.g., rest tremor, rigidity, and bradykinesia). In addition to traditional clinical symptoms, neuroimaging, genetic, and biochemical studies are applied for early detection of PD, whereas no reliable biomarkers can be used as the only valid criterion. Consequently, patients that are underdiagnosed and misdiagnosed are common [2]. A single biomarker is insufficient for the PD diagnosis, and a variety of methods need to be combined.

Speech characteristics that can serve as a novel clinical biomarker for PD diagnosis have been noticed by researchers. A longitudinal study has shown that the disturbance of speech acoustic parameters in Parkinson's disease (atypical range and variance of the fundamental frequency) occurs approximately 5 years before the onset of clinical symptoms [3]. Other related studies mainly focus on the speech differences between PD and healthy speakers which manifest in multiple speech subsystems, including phonation, articulation, and prosody, which can be used to distinguish PD patients from healthy adults [4].

Although speech disorder may be an early sign of PD, it is usually ignored by patients themselves and their caregivers [5, 6]. So, investigating the speech characteristics in PD patients, especially for patients in their early stages, is vital to early diagnosis of PD patients. Nevertheless, few studies have reported the contributions of speech features in the identification of PD patients.

## 2. Literature Review

As stated by previous studies, changes in prosodic and segmental speech features of speech (e.g., speech rate and articulatory deficits) are among the early symptoms in PD

patients, whereas the difference in phonation features of Parkinsonian and healthy speech is not significant [5,7]. So, prosodic and segmental features of PD speech are reviewed in the current study.

Less variability of fundamental frequency (F0) and narrower F0 range which are in close relation with monopitch or monotone is the most prominent feature of PD speech [8]. Previous studies have confirmed that F0 disorders in PD speakers exist not only in the prodromal stage but also in the later stage of the disease [7]. Ruzs et al. [9] reported that F0 variability is one of the most reliable acoustic indicators of Parkinson's disease. Studies targeting the relationship between acoustic features of PD speech and their motor symptoms show a significant negative correlation between F0 variability of parkinsonian speech and disease progression, with an explanation that the movement of vocal folds in PD patients is sensitive to reflect disease progression [10].

Deficits in speech timing are also observed in early PD speakers [11, 12]. Speech rate is affected by Parkinson's disease [13]. Compared to healthy speakers, PD patients speak at a faster or slower rate, showing a great individual variability [14]. Although speech rate has been used to study the pathological change in PD speech, its reliability to differentiate speakers with PD from healthy speakers remains uncertain. Rhythmic metrics play an important role in distinguishing pathological speech from normal speech. Liss et al. [15] stated for the first time that a set of rhythmic metrics (e.g., standard deviation of consonantal durations and proportion of vocalic durations) can be used to distinguish speakers with dysarthria from healthy individuals. In particular, Lowit et al. [12] found the proportion of vocalic durations in speech is a robust indicator of PD.

Studies of motor speech disorders have been focused on vowel production to investigate articulation impairments, and thus articulation deficits have been extensively studied in PD speakers. Vowel articulation impairment is considered an important marker for early-stage PD speakers [16, 17]. Various indices such as the ratio between the second formant of the vowels /i/ and /u/ ( $F2_i/F2_u$ ), vowel space area (VSA), and vowel articulation index (VAI) are used to characterize vowel articulation in previous studies; however, these studies show inconsistent results. Several studies show that VSA is less reliable than VAI in identifying PD speakers suffering from mild dysarthria [17–19], while other studies conclude that both VAI and VSA can identify articulation disorders of PD speakers [16].

Inconsistent results are also observed for stops in the PD speech. Longer voice onset times (VOTs) of consonants have been observed in PD speech [20, 21]. On the other hand, Ackermann and Ziegler [22] and Kim [23] found that stops have shorter VOTs in PD speech. Moreover, a recent study has found that there is no significant difference in VOT between speakers with PD and healthy individuals [24].

Although acoustic differences in speech between individuals with PD and healthy speakers are well documented, little is known about which acoustic features are the most important in distinguishing PD patients, in particular early-stage ones, from healthy adults. In contrast, neurologists

usually use clinical scales such as the United Parkinson's Disease Rating Scale [25], which however is subjective and may mislead their judgments of PD patients.

Therefore, this study focuses on two questions. First, which metrics are the most important in differentiating PD from healthy speakers? Second, is automatic classification more advantageous in identifying PD patients than auditory perception? To investigate these two questions, the machine learning algorithm based on both prosodic and segmental features of speech is used to explore the relative contributions of acoustic features and the accuracy rate of early-stage PD identification. Then, neurologists are recruited to judge whether the speaker is a PD patient or not after they hear the reading speech. The accuracy rate is compared with that of automatic classification.

### 3. Methods

To solve the two problems raised above, this study recruited PD patients in early stages and recorded their reading speech in quiet rooms (noise < 50 dB). Based on the reading speech, several acoustic metrics of speech are extracted. All speech data are divided into two sets, i.e., training set and testing set. Using random forest classification, we explored the accuracy rate of classification of early-stage PD patients and the relative contributions of these acoustic metrics to the classification. The framework of this study is shown in Figure 1.

**3.1. Participants and Materials.** To minimize the effects of dialects and special speaking styles on acoustic results, speakers who spoke with an accent or spoke in an unusual way were screened out [26]. So, thirty-six individuals with idiopathic PD (19 men and 17 women) aged 52 to 78 years (mean = 63.55, SD = 9.46) were recruited as the PD group. The Hoehn and Yahr score of PD was between 1 and 2.5, and the disease duration was less than 5 years, which meant that all patients were in the early stage of PD [27]. None of the patients had suffered from other diseases or had undergone deep brain stimulation surgery. All participants were native Mandarin speakers and scored at least 24 on the Mini-Mental State Exam.

Speech task is an important factor in voice disorder investigation [11, 28]. According to previous studies, passage reading is the optimum task to explore PD speech compared with sustained vowels and fast syllable repetition [21, 29], so passage reading is used in the current study. Before speech recording, the participants were instructed to read silently the passage *The North Wind and the Sun*, which consisted of 169 syllables in Mandarin. Then, all participants read the passage in a quiet room and speech was recorded using a Zoom H4n portable recorder at a sampling rate of 44100 Hz. The English translation of this passage is shown below.

The North Wind and the Sun were disputing which was the strongest, when a traveller came along wrapped in a warm cloak. They agreed that the one who first succeeded in making the traveller take his cloak off should be considered stronger than the other. Then, the North Wind blew as hard as he could, but the more he blew, the more closely did the

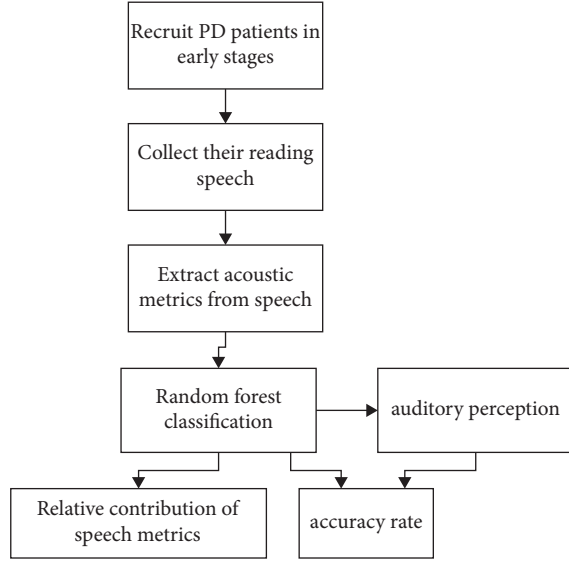


FIGURE 1: Framework of this study.

traveller fold his cloak around him, and at last the North Wind gave up the attempt. Then, the Sun shined out warmly, and immediately the traveller took off his cloak, and so the North Wind was obliged to confess that the Sun was the strongest of the two.

To avoid the influence of anti-Parkinson medication on speech production, speakers of PD stopped taking medication and were in a fasting state for at least 12 hours before the speech recording.

**3.2. Acoustic Measures.** Acoustic analysis was conducted on two aspects of acoustic measures. One aspect was the prosodic feature (fundamental frequency parameters, speech rate, and rhythm), and the other was the segmental feature (formants of vowels and voice onset time of stops). Details of the 13 acoustic measures are described below.

The measures of the fundamental frequency ( $F0$ ) were extracted from the Praat software. The autocorrelation algorithm in Praat generated  $F0$  tracks for every speech sample, and then gross  $F0$  errors were manually corrected according to the waveform and spectrogram. Five  $F0$  parameters were calculated from the  $F0$  tracks, namely, the minimum ( $F0_{\min}$ ), maximum ( $F0_{\max}$ ), mean ( $F0_{\text{mean}}$ ),

range (difference between  $F0_{\min}$  and  $F0_{\max}$ ,  $F0_{\text{range}}$ ), and standard deviation of  $F0$  ( $F0_{\text{std}}$ ). All  $F0$  values were converted from Hz to semitones (St) with 50 Hz as the reference frequency. The formula is as follows (fr is the  $F0$  value in Hz).

$$\text{St} = 12 \times \log_2\left(\frac{\text{fr}}{50}\right). \quad (1)$$

Both speech rate and articulation rate were calculated. Speech rate was defined as the total number of syllables divided by the total speech duration, and articulation rate was defined as the total syllables divided by the articulation duration. Pauses longer than 200 ms were marked and excluded from the measurement of articulation rate [30].

$$\text{Speech rate} = \frac{\text{the numbers of syllables}}{\text{total speech duration}},$$

$$\text{articulation rate} = \frac{\text{the number of syllables}}{\text{total speech duration} - \text{pause duration}}. \quad (2)$$

Among the rhythmic parameters commonly used to distinguish different types of language [31], proportion of vowel vocalic durations in speech (%V) was reported to be effective in identifying speakers with brain injury and Parkinson's disease as reported [12, 32], so %V was investigated as a rhythmic measure in this research.

$$\%V = \frac{\text{vocalic duration}}{\text{total speech duration}} \times 100\%. \quad (3)$$

Vowel articulation was measured by the ratio of the second formants of the vowels /i/ and /u/ ( $F2i/F2u$ ), and the triangular vowel space area (tVSA) and the vowel articulation index (VAI) were based on the corner vowels /a, i, u/ in the speech sample. The middle 60% interval of the whole vowel was extracted from the monosyllables containing these three vowels to obtain the stable values of the first and second formants ( $F1$  and  $F2$ ) of these three vowels. tVSA with unit  $\text{Hz}^2$  has been widely used to measure the distribution of vowels, and VAI has been associated with the concentration or dispersion of vowels. The more concentrated the vowels, the smaller the VAI (the minimum of VAI is 0.5). The formulas for tVSA and VAI are as follows [17]:

$$\text{tVSA} = \text{ABS} \left\{ \frac{[F1[i] \times (F2[a] - F2[u]) + F1[a] \times (F2[u] - F2[i]) + F1[u] \times (F2[i] - F2[a])]}{2} \right\}, \quad (4)$$

$$\text{VAI} = \frac{F2[i] + F1[a]}{F1[i] + F1[u] + F2[u] + F2[a]}.$$

Voice onset time (VOT) of stops reflects the precise time coordination between the movements of the supra laryngeal articulators (such as lips and tongue) and the vocal folds. In this study, the VOTs of the post-pausal stops

/p, p<sup>h</sup>, t, t<sup>h</sup>, k/ were compared between the two groups. Considering that speech rate has a direct impact on VOT, a normalized parameter, VOT ratio, was used in this study [24].



$$\text{VOT ratio} = \frac{\text{VOT}}{\text{syllable duration}} \times 100\%. \quad (5)$$

Because VOT is longer for aspirated stops than for unaspirated stops in Mandarin, the VOT ratios of post-pausal stops were calculated for aspirated  $/p^h, t^h/$  and unaspirated  $/p, t, k/$  (hence VOT ratio\_un and VOT ratio\_as), respectively. All the 13 acoustic measures are summarized in Table 1.

**3.3. Random Forest Classification.** Based on all these acoustic parameters, a random forest classification in R [33] was used to investigate the contribution of these acoustic parameters to the classification of PD individuals and healthy controls.

First, the minimum value of the mean error rate of the model was calculated based on out-of-bag (OOB) data, and the optimal number of variables of the binary tree in the node “mtry” was set to 3 in the package randomForest. Since the model error tends to converge when the number of decision trees approaches 500, the optimal number of decision trees nTree was set to 500. Among the experimental data, 70% were randomly selected as the training set and the remaining 30% were selected as the test set. Finally, the accuracy, sensitivity, and specificity were calculated.

**3.4. Auditory Perception Test.** To compare the effectiveness of acoustic classification and the auditory judgment by neurologists, an auditory perception test was performed.

An utterance of the following sentence for about 10 seconds: “They agreed that the one who first succeeded in making the traveller take his cloak off should be considered stronger than the other,” was selected from speech recording from each participant, so that 36 stimuli were collected as auditory stimuli. Five neurologists who had been working in the department of neurology for at least 2 years were recruited in the auditory perception test. All five listeners had rich clinical experience in the assessment of patients with PD, but they were unaware of the purpose of the study.

After the intensity was normalized to 70 dB using the Praat software, all 36 speech stimuli were played back in a random order to the listeners through earphones in a quiet room. The following prompt was displayed on the computer screen: “Please judge from the speech you heard whether the speaker is a patient with PD.” Then, after each speech stimulus was played back, two options appeared on the computer screen: “PD patients” and “healthy persons.” Listeners clicked the mouse to complete the forced choice.

## 4. Results

**4.1. Results of Random Forest Classification.** The results based on 13 acoustic measures showed that the classification accuracy of the random forest algorithm was 75.6%, and the sensitivity and specificity were 66.7% and 84.6%, respectively.

Figure 2 shows the importance of all 13 acoustic measures in distinguishing PD patients from healthy controls.

TABLE 1: Acoustic measures extracted from speech.

Acoustic measure	Type of feature
F0min (St)	Prosody
F0max (St)	
F0mean (St)	
F0range (St)	
F0std (St)	
Articulation rate (syl/s)	
Speech rate (syl/s)	
%V	Segments
F2i/F2u	
tVSA (kHz <sup>2</sup> )	
VAI	
VOT ratio_un	
VOT ratio_as	

The most important measures in identifying speakers with early PD were F0std, F2i/F2u, F0 range, and VAI.

**4.2. Results of Auditory Perception Test.** The perceptual accuracy of five neurologists ranged from 61.1% to 68.1%, with a mean accuracy of 64.2%, which was lower than the accuracy of random forest classification (75.6%).

## 5. Discussion

Speech disorders have been well documented in previous studies, whereas the current study focuses on the contributions of prosodic and segmental features of early PD speakers and the superiority of acoustic classification using the random forest algorithm to clinical auditory judgments.

The random forest classification based on 13 acoustic measures including both prosodic and segmental features showed that the most important measures in identifying early PD speakers were F0std, F2i/F2u, F0 range, and VAI. Among these 4 acoustic measures, F0std and F0 range were prosodic features while F2i/F2u and VAI were segmental features. The smaller F0 range and F0 variability (F0std) in the early-stage PD speakers suggested their less expressive pitch variation, consequently resulting in monotone as previous studies reported [3, 9]. Although previous studies reported that either F0 parameters or vowel articulation played roles in distinguishing PD from healthy speakers [9, 17], the current study further showed that both F0 and vowel articulation were important in identifying individuals with early PD.

The identification accuracy of speakers with early-stage PD was 75.6%, which was higher than that of auditory assessment by neurologists. Therefore, this objective algorithm could be an effective auxiliary method for neurologists to identify speech disorders in early-stage patients with PD that were not easily perceptible.

It should be noted that the accuracy of identifying speakers with early PD based only on acoustic measures was 75.6%, still far away from being perfect. There were two possible reasons. First, not all speakers with early PD had speech disorders. Second, in addition to these acoustic measures we investigated here, some other parameters may

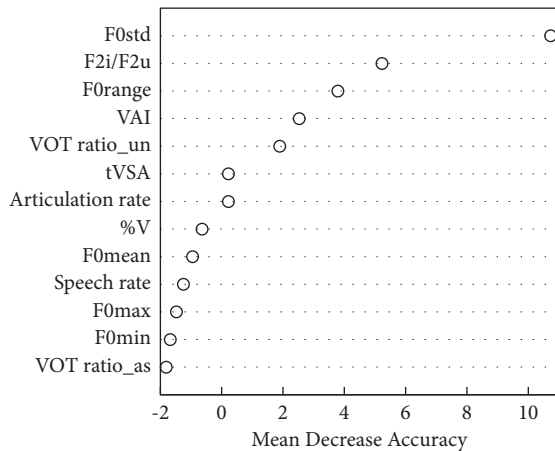


FIGURE 2: Importance of acoustic measures.

also play roles in distinguishing early PD patients from healthy individuals. Therefore, our future study will explore more acoustic metrics and conduct multiple tasks, e.g., monologue and dialogue, to make a better identification of early-stage PD.

## 6. Conclusion

Random forest classification algorithm based on acoustic measures, especially fundamental frequency and vowel articulation, can be an auxiliary method to identify early-stage PD. There are diverse early biomarkers of PD such as clinical symptoms and neuroimages. However, the value of a single biomarker for early diagnosis is limited, and a variety of methods can be combined to improve the accuracy of diagnosis. The result of our research shows that a classification model based on speech acoustic parameters can provide a more economical, convenient, and effective way for early diagnosis, accurate assessment, and remote monitoring of PD patients. Nevertheless, acoustic parameters extracted from reading speech in this study are not yet comprehensive, and more acoustic features and multiple speech tasks should be considered in future studies.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The author declares that there are no conflicts of interest.

## Acknowledgments

This research was supported by the Major Program of the National Social Science Fund of China (13&ZD189). The author would like to thank Prof. Wentao Gu for his supervision and Dr. Weiguo Liu for his help in recruiting the PD patients.

## References

- [1] M. S. Islam, K. A. Mamun, and D. Hai, "Decoding of human movements based on deep brain local field potentials using ensemble neural networks," *Computational Intelligence and Neuroscience*, vol. 2017, p. 16, Article ID 5151895, 2017.
- [2] A. J. Hughes, S. E. Daniel, S. Blankson, and A. J. Lees, "A clinicopathologic study of 100 cases of Parkinson's disease," *Archives of Neurology*, vol. 50, no. 2, pp. 140–148, 1993.
- [3] B. Harel, M. Cannizzaro, and P. J. Snyder, "Variability in fundamental frequency during speech in prodromal and incipient Parkinson's disease: a longitudinal case study," *Brain and Cognition*, vol. 56, no. 1, pp. 24–29, 2004.
- [4] S. Sapir, "Multiple factors are involved in the dysarthria associated with Parkinson's disease: a review with implications for clinical practice and research," *Journal of Speech, Language, and Hearing Research*, vol. 57, no. 4, pp. 1330–1343, 2014.
- [5] J. Hlavnička, R. Čmejla, T. Tykalová, K. Šonka, E. Růžička, and J. Ruzs, "Automated analysis of connected speech reveals early biomarkers of Parkinson's disease in patients with rapid eye movement sleep behaviour disorder," *Scientific Reports*, vol. 7, no. 1, p. 12, 2017.
- [6] R. B. Postuma, A. E. Lang, J. F. Gagnon, A. Pelletier, and J. Y. Montplaisir, "How does parkinsonism start? Prodromal parkinsonism motor changes in idiopathic REM sleep behaviour disorder," *Brain*, vol. 135, no. 6, pp. 1860–1870, 2012.
- [7] J. Ruzs, J. Hlavnička, T. Tykalová et al., "Quantitative assessment of motor speech abnormalities in idiopathic rapid eye movement sleep behaviour disorder," *Sleep Medicine*, vol. 19, pp. 141–147, 2016.
- [8] L. K. Bowen, G. L. Hands, S. Pradhan, and C. E. Stepp, "Effects of Parkinson's disease on fundamental frequency variability in running speech," *Journal of Medical Speech-Language Pathology*, vol. 21, no. 3, pp. 235–244, 2013.
- [9] J. Ruzs, R. Čmejla, H. Ruzickova, and E. Ruzicka, "Quantitative acoustic measurements for characterization of speech and voice disorders in early untreated Parkinson's disease," *Journal of the Acoustical Society of America*, vol. 129, no. 1, pp. 350–367, 2011.
- [10] F. Majdinasab, S. Karkheiran, M. Soltani, N. Moradi, and G. Shahidi, "Relationship between voice and motor disabilities of Parkinson's disease," *Journal of Voice*, vol. 30, no. 6, pp. e17–e22, Article ID 768.
- [11] Z. Galaz, J. Mekyska, Z. Mzourek et al., "Prosodic analysis of neutral, stress-modified and rhymed speech in patients with Parkinson's disease," *Computer Methods and Programs in Biomedicine*, vol. 127, pp. 301–317, 2016.
- [12] A. Lowit, A. Marchetti, S. Corson, and A. Kuschmann, "Rhythmic performance in hypokinetic dysarthria: relationship between reading, spontaneous speech and diadochokinetic tasks," *Journal of Communication Disorders*, vol. 72, pp. 26–39, 2018.
- [13] Y.-F. Chiu, A. Neel, and T. Loux, "Acoustic characteristics in relation to intelligibility reduction in noise for speakers with Parkinson's disease," *Clinical Linguistics and Phonetics*, vol. 35, no. 3, pp. 222–236, 2020.
- [14] A. M. Goberman and C. Coelho, "Acoustic analysis of parkinsonian speech I: speech characteristics and L-dopa therapy," *NeuroRehabilitation*, vol. 17, no. 3, pp. 237–246, 2002.
- [15] J. M. Liss, L. White, S. L. Mattys et al., "Quantifying speech rhythm abnormalities in the dysarthrias," *Journal of Speech, Language, and Hearing Research*, vol. 52, no. 5, pp. 1334–1352, 2009.



## Research Article

# Automatic Rule Generation for Decision-Making in Context-Aware Systems Using Machine Learning

Roua Jabla <sup>1,2</sup> Maha Khemaja <sup>3</sup> Félix Buendia <sup>1</sup> and Sami Faiz <sup>4</sup>

<sup>1</sup>Department of Computer Engineering, Universitat Politècnica Valencia, Camino de Vera S/N, Valencia 46022, Spain

<sup>2</sup>ISITCom, University of Sousse, Sousse 4011, Tunisia

<sup>3</sup>PRINCE Research Lab, ISITCom, University of Sousse, Sousse 4011, Tunisia

<sup>4</sup>LTSIRS Laboratory, University of Tunis El Manar, Tunis 5020, Tunisia

Correspondence should be addressed to Roua Jabla; [jabla.roua@gmail.com](mailto:jabla.roua@gmail.com)

Received 11 December 2021; Revised 13 March 2022; Accepted 15 April 2022; Published 6 May 2022

Academic Editor: Daqing Gong

Copyright © 2022 Roua Jabla et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the increasing interest devoted to dynamic environments, a crucial aspect is revealed in context-aware systems to deal with the rapid changes occurring in users' surrounding environments at runtime. However, most context-aware systems with predefined context-aware rules may not support effective decision-making in dynamic environments. These context-aware rules, which take into account different context information to reach an appropriate decision, could lose their efficiency at runtime. Therefore, a growing need is emerging to address the decision-making issue leveraged by dynamic environments. To tackle this issue, we present an approach that relies on improving decision-making in the wake of dynamic environments through automatically enriching a rule knowledge base with new context-aware rules discovered at runtime. The major features of the presented approach are as follows: (i) a hybridization of two machine learning algorithms for rule generation, (ii) an extended genetic algorithm (GA) for rule optimization, and (iii) a rule transformation for the knowledge base enrichment in an automated manner. Furthermore, extensive experiments on different datasets are performed to assess the effectiveness of the presented approach. The obtained experimental results depict that this approach exhibits better effectiveness compared to some algorithms and state-of-the-art works.

## 1. Introduction

In context-aware computing, there exists a shift from static environments to dynamic environments [1]. This shift reflects a growing interest devoted to dynamic environments. With the growing interest, a crucial need is revealed for context-aware systems to be aware of and to adapt to their changing contexts in highly dynamic environments at runtime [2]. To support this need, a grand challenge is that context-aware systems should adjust their behaviors to the dynamics entailed in their surrounding environments at runtime. In order to meet this challenge, a decision-making process improvement by providing appropriate services to users situated in highly changing environments has emerged to make context-aware systems more resilient to dynamic environments at runtime. Nevertheless, at present, most

context-aware systems usually work well in static environments with predefined context-aware rules that can only behave to changes in environment attributes and context information [3]. They can handle neither dynamic environments nor context changes at runtime since predefined rules might not be suitable for the dynamic nature of environments. On that account, rules need to be constantly evolved to remain relevant over time [4]. This raises more attention to be paid to how to enrich a rule knowledge base through the generation of efficient rules to timely react to arising environment changes at runtime. As data mining and specifically machine learning could be more accurate for profiting better rules [5], exploring recent advances in the application of machine learning might be beneficial in offering a way to learn and generate a set of nonredundant rules that are easily comprehensible and are capable of

representing contextual knowledge in a very clear and efficient way.

To achieve this sustained attention, we propose, in this work, an approach that aims to automatically improve the decision-making process to support dynamic environments at runtime. The main feature of the proposed approach is to offer a rule knowledge base, where context-aware rules are fluid and evolutive at runtime for alleviating the burden of manually creating rules to react toward users' environment changes. The novelty and contribution of this approach could be drawn from threefold: first, we present a novel hybrid learning approach toward effectively generating a concise set of nonredundant association rules in an automated fashion by applying two learning algorithms. We hybridize machine learning algorithms to generate a more accurate and complete set of rules, to avoid redundancy, and to build a strong rule knowledge base in a context-aware system, in which we are interested. Second, we extend a genetic algorithm (GA) [6] with a multianalysis technique in the direction of rule optimization. The rule optimization is applied to find the well-performed rules from the generated rules since machine learning algorithms are not much proficient at optimizing rules. Third, we introduce an automatic transformation of obtained association rules to rules expressed in Jena [7]. Rule transformation is performed to express the well-performed rules in such a way that they can run over an ontology-based context model and a context-aware system can reason. Moreover, we conduct a range of experiments to assess the effectiveness of the proposed approach on different datasets from the UCI Machine Learning Repository [8]. First, we compare the proposed approach with the traditional association rule mining algorithms to analyze the number of generated rules. Then, we evaluate the performance of the proposed approach with the most common machine learning algorithms and certain state-of-the-art works. Finally, we analyze the computational time of the proposed approach and the machine learning algorithms. The provided results prove the effectiveness of the proposed approach by achieving the best result in terms of the number of rules, precision, recall, and accuracy among certain state-of-the-art works and algorithms such as apriori, FP growth, K-nearest neighbor, Naïve Bayes, JRip, and decision table.

The rest of this study is arranged as follows. We review the related work in Section 2. In Section 3, we introduce the proposed approach and outline the overall architecture and modules in detail. We present the experimental setup and the results obtained with an example of generated context-aware rules in Section 4. In Section 5, we illustrate the discussion of the results. Finally, we draw conclusions and highlight the direction for future work in Section 6.

## 2. Related Work

Data mining, especially machine learning and association rule mining, algorithms play a vital role in rule discovery from data. By searching through the literature, there is an extensive research basis to support association rule discovery from data to tackle the challenge of the decision-making

improvement. In the following, we review first some works for discovering rules within the area of context awareness. Then, we provide a discussion to highlight the research gaps that motivate us to propose our approach.

A large number of works for association rule learning have been introduced using association rule mining algorithms. In this sense, Gabroveanu and Diaconescu [9] proposed a recommender system for students. They used data obtained from the learning database and apriori as an association rule mining algorithm in order to identify strong association rules for students. Then, these rules, obtained in an offline mining process, are translated into Jena rules to allow reasoning over RDF models. In addition, Kaliappan and Sai [10] presented a new modified apriori algorithm for finding the association rules among large datasets to promote sales and user interaction. They showed that the proposed algorithm improved the efficiency of generating association rules. Moreover, Davagdorj and Ryu [11] offered an association rule mining method to discover useful patterns, which include medical knowledge, from a medical dataset. They applied the FP growth algorithm to extract a set of association rules. Then, the obtained rules are used to support medical decision-making for interpreting diagnosing patient information. Furthermore, Asadianfam et al. [12] introduced a new approach to improve recommendations that can be used to predict the next navigable page of users. One of the objectives considered in their approach is to provide appropriate recommendations to users who have different profiles from the existing users' profiles. To deal with the objective, the authors used the apriori algorithm to generate association rules from users' behaviors and then made appropriate recommendations. They showed that the generated association rules could increase the overall efficiency of the recommender system. More recently, Miswan et al. [13] proposed a framework of association rule mining in readmission tasks. The proposed framework consisted of two processes, namely, data preprocessing and rule mining extraction. Apriori algorithm is used to extract the hidden input variable patterns and relationships among admitted patients by generating supervised learning rules. The mined rules are discussed and validated by the domain expert, which is a valuable guide in making decisions on targeted patients' clinical resources based on various readmission durations.

Apart from these works, there are also few works exploring the association rule learning using machine learning algorithms. In this context, Hong et al. [14] proposed an agent-based framework for offering personalized services utilizing the extracting users' preferences and association rules. The decision tree algorithm is considered to infer association rules for recommending personalized services for users. Similarly, Zulkernain et al. [15] introduced an intelligent mobile interruption management system. The main idea of their proposed system is to intelligently assist users in their daily activities. To this end, a decision tree algorithm is used to make intelligent decisions. Moreover, Sarker [16] presented an association rule learning approach that can be used to discover a set of nonredundant and useful rules. In their approach, they considered, first, the Naïve

Bayes (NB) algorithm to eliminate noise from data and, second, the decision tree algorithm to generate a set of association rules. These algorithms are used to build a robust prediction model that could improve the prediction accuracy. Finally, Basha [17] provided a cardiovascular prediction system that combines the traditional K-nearest neighbor (KNN) algorithm with a GA to extract strong association rules for facilitating the decision-making process. First, the proposed system extracts association rules using the KNN algorithm. Then, the output rules become the population of the GA to remove redundant and irrelevant rules.

Nevertheless, a common weakness that can be found in the majority of discussed works [9–13] mainly stands in the use of traditional association rule mining algorithms, such as FP growth and apriori. This weakness is caused by the narrow applicability of these algorithms due to the huge number of association rules generated [18]. However, these works generate numerous redundant association rules, which lead to generating a huge number of uninteresting rules that become useless in making decisions. This redundant generation makes not only the rule set unnecessarily large but also makes the decision-making process more complex and ineffective. Therefore, the traditional association rule mining algorithms are not able to efficiently extract interesting association rules. Instead, employing machine learning algorithms can overcome this weakness and avoid rule redundancy and conflicts in the mining process. Despite that, in many cases, works based on machine learning algorithms [14–17] could not ensure high accuracy in generating association rules [19]. Furthermore, a limitation in some of these works [9, 13] is that they are only suitable for offline mining and not suitable for dynamic environments where new incoming changes are continuously received at runtime and so decisions may be wrongly predicted based on rules defined a priori. A further important limitation is that certain works like the work of Miswan et al. [13] require domain expert intervention to validate generated association rules. Following the above discussion, we aim to close gaps within the discussed works by proposing an approach that relies on automatically enriching a rule knowledge base with new context-aware rules generated at runtime to perform an efficient decision-making for dynamic environments. The proposed approach focuses on discovering and generating nonredundant and well-performed association rules through the use of machine learning algorithms along with an extended GA to reduce rule redundancy. To address the decision-making accuracy question, the proposed approach represents a hybridization idea that combines the strengths of two supervised machine learning algorithms, the decision tree, and the random tree algorithms. Moreover, it proceeds for a rule transformation to support the automatic enrichment of rule knowledge bases at runtime.

### 3. Proposed Approach

In this work, we propose an approach that supports the improvement of decision-making in dynamic environments

at runtime by allowing an automatic enrichment of a rule knowledge base with new generated context-aware rules. The proposed approach first aims to generate a concise set of nonredundant association rules following the IF-THEN structure, then to optimize the generated rules, and finally to transform rules to Jena rules for enriching the rule knowledge base and providing appropriate services to users situated in dynamic environments at runtime. Figure 1 illustrates a schematic view of the proposed approach architecture.

As illustrated in Figure 1, the proposed approach typically consists of two main modules, namely, the rule generation module and the rule transformation module. In the following, we discuss these modules and their roles in generating association rules and transforming them into Jena rules.

*3.1. Rule Generation Module.* The rule generation module is designed to automatically derive association rules needed to meet the changes occurring in users' dynamic environments. That means, it aims to preprocess a candidate dataset, to generate decision trees, and to infer the well-performed association rules in order to further enrich a rule knowledge base and improve the decision-making process at runtime. The present module runs every time when new changes are arrived and a priori rules are deemed not relevant to these changes. This module is defined considering two main phases as shown in Figure 1.

*3.1.1. Association Rule Learning.* The association rule learning phase is in charge of learning and extracting association rules from a candidate dataset. This phase includes various steps starting from the preprocessing of the dataset to the validation of rules.

*(1) Dataset Preprocessing.* This represents the first step of the association rule learning phase, which is applied to improve the quality of the candidate dataset so as to ensure accurate, consistent, and complete generated decision trees and further association rules. This step is responsible for preparing data from the candidate dataset. Hence, two major stages during the data preprocessing step, namely, data cleaning and data reduction, are performed. The first stage comprises both operations like filling missing data and smoothing noisy data. It aims to replace missing data with the average of existing data, using different traditional imputation methods (i.e., mean and median). To smooth noisy data, we carry out techniques, such as clustering, regression, and binning, to eliminate the noise in the dataset as it rises due to random variation. The second stage is data reduction that is used to obtain a simplified representation of a dataset with relevant data. To do it in the simplest manner, we remove redundant and inconsistent data.

*(2) Learning-Based Rule Generation.* After the data preprocessing step is completed, the learning-based rule generation step takes place first to automatically learn training models for building decision trees, then to derive association

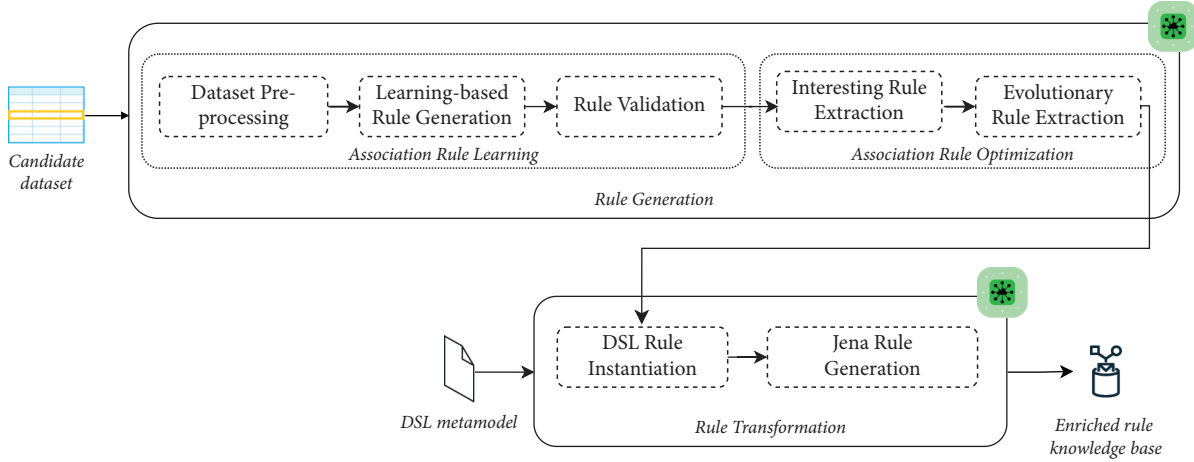


FIGURE 1: Proposed approach architecture.

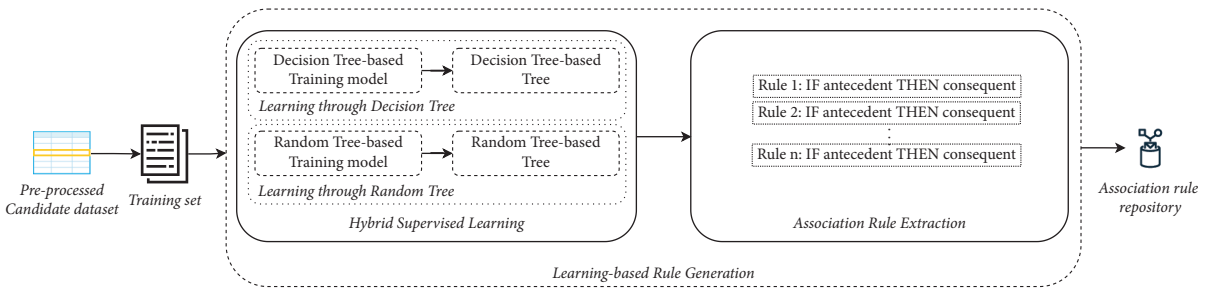


FIGURE 2: Learning-based rule generation step.

rules, and finally to integrate them into the association rule repository. To this end, the present step provides two elementary mechanisms. The main idea of these mechanisms can be simply described in Figure 2.

A first mechanism, called hybrid supervised learning, takes advantage of machine learning and pattern recognition in learning about data, extracting relevant relationships, and eliminating redundancy. Thus, it combines widely adopted supervised machine learning algorithms, namely, J48 decision tree [20] and random tree [21], to train the preprocessed dataset. As far as we know, this is the first time that J48 decision tree and random tree algorithms are combined to the association rule learning. To improve the accuracy of learning, we come up with the idea of hybridizing J48 decision tree and random tree algorithms as they are tree-based techniques that provide the highest accuracy diagram of a decision tree [22]. The resulting training models consist of a set of decisions in a tree structure, which could be utilized to generate rules from each leaf node [23]. As shown in Algorithm 1, this first mechanism trains the J48 decision tree to build the corresponding decision tree on the decision tree-based training model. Subsequently, it applies the random tree algorithm to get the second decision tree on the random tree-based training model.

The second mechanism, called association rule extraction, acquires association rules from previously trained decision trees. To do so, it tracks, in an automated fashion,

the path from the root node to each leaf node in both trees in order to detach the set of association rules.

The output of this mechanism is a set of association rules following IF-THEN statements: IF  $\langle A \rangle$  THEN  $\langle C \rangle$ , where the antecedent part  $\langle A \rangle$  represents user's surrounding contextual information such as temporal context, spatial context, social contexts, or others relevant contextual information, for example, "outlook = overcast," and the consequent part  $\langle C \rangle$  represents their corresponding behavioral activities for decision-making, for example, "play = yes."

(3) *Rule Validation*. After completing the generation step, the rule validation step is applied through the following methods:

- (i) Rule structure verification, which is responsible for checking whether extracted rules are preserving the inherent IF-THEN structure.
- (ii) Rule consistency verification, which is in charge of verifying the consistency of association rules. Given the fact that association rules are made up of antecedent constraints and a consequent constraint, the rule consistency is related to the satisfiability of constraints. To ensure a conflict-free association rule repository, a reasoner is used to enumerate all inconsistent rules, where the consequent constraint does not refer to the antecedent constraint.



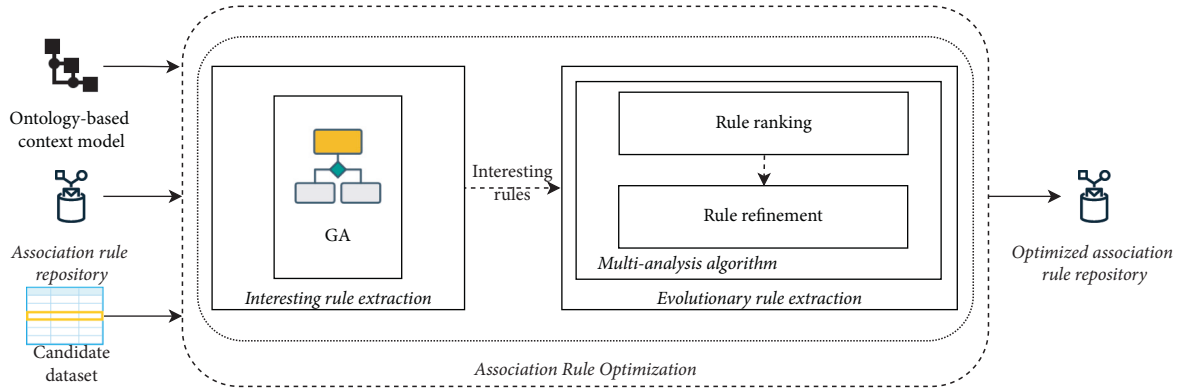


FIGURE 3: Rule optimization phase.

**3.1.2. Association Rule Optimization.** The association rule optimization is the second and last phase in the rule generation module. The present phase is in charge of identifying the well-performed rules from a set of earlier validated association rules. In this sense, a GA is extended to support a multianalysis technique. The main idea of extending a GA with a multianalysis technique is to optimize the set of association rules to find the well-performed rules by adopting the optimization strategies in GA. To be more precise, there are two main steps in this phase as depicted in Figure 3. First, the interesting rule extraction step is carried out using GA that is a kind of effective optimizing technique owing to the robust and global search ability [24]. Second, the evolutionary rule extraction step involves a multianalysis technique to extend the GA and extract the well-performed association rules.

*(1) Interesting Rule Extraction.* This step determines the interesting association rules by applying a GA to the candidate dataset. The GA is used to produce a strong level of association rules since supervised learning algorithms may generate irrelevant rules. The motivation behind this choice is twofold: first, the GA is one of the best methods for rule optimization [25], and second, it performs a global search technique to find out interesting rules with less complexity compared to other algorithms [26]. The GA creates an initial population as a collection of chromosomes, in which every chromosome represents an association rule. Then, it evolves the initial population over multiple generations through encoding, selection, crossover, and mutation operations to reach the optimal set of interesting association rules. In the end, it introduces a set of interesting association rules, which satisfies a fitness function. The flowchart of GA is shown in Figure 4.

The specific steps of the GA are as follows:

- (i) Step 1. Encoding step: The candidate dataset is encoded to initiate the experimentation of the GA. In our case, a binary encoding schema is used.
- (ii) Step 2. Initial population generation: An initial population of size  $K$  chromosomes is randomly generated as a set of solutions to be optimized. These chromosomes are a representation of the rules generated from the candidate dataset.

- (iii) Step 3. Calculating fitness: The fitness value of each chromosome in the population is calculated by a fitness function to find the association rules that their support and confidence are larger than other rules. To this end, a fitness function that has been described in the work of Qodmanan et al. [27] is considered as given in equation.

$$\text{fitness} = \frac{(1 + \text{supp}F(A \cup C))^2}{1 + \text{supp}(A)}. \quad (1)$$

In equation (1),  $\text{supp}(A \cup C)$  is the support of the  $A \rightarrow C$  and  $\text{supp}(A)$  is the support of the antecedent part of it.

- (iv) Step 4. Selection: A chromosome with a high fitness value is selected from the population on the basis of a fitness function.
- (v) Step 5. Crossover: The next generation of the population is generated based on the calculated fitness values. The idea behind crossover is to combine the two parent chromosomes to produce two new offspring. The result of crossover is the birth of two new chromosomes. Crossover is carried out according to a defined crossover probability.
- (vi) Step 6. Mutation: Mutation randomly changes chosen bits from 0 to 1 or from 1 to 0. It is applied to the new offspring with a certain mutation probability. The purpose is to maintain diversity among the different generations to increase the global optimization of the GA.
- (vii) Step 7: After a series of selection, crossover, and mutation, the GA is stopped when the generated chromosomes meet the optimality or the maximum number of generations. Otherwise, it turns back to step 3 to continue the rule optimization.

*(2) Evolutionary Rule Extraction.* In the previous step, the interesting association rules are selected. However, we cannot get the well-performed rules that could lead to achieve the appropriate decision-making performance due to the fact that the fitness function, defined in GA, might be

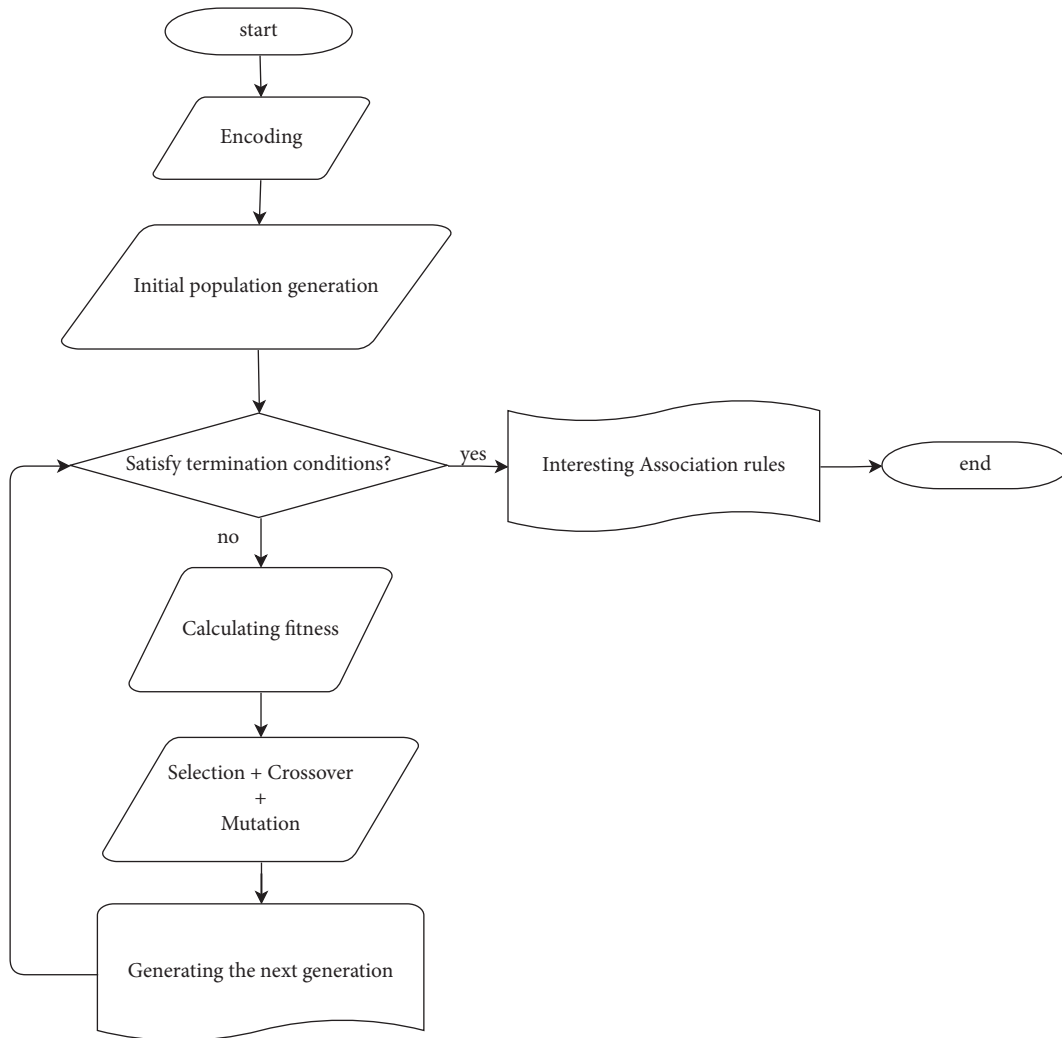


FIGURE 4: Flowchart of a GA.

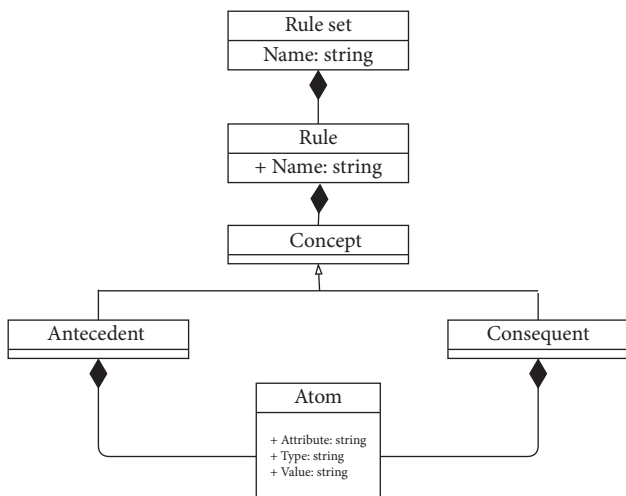


FIGURE 5: Jena rule definition metamodel.

in conflict. To deal with this issue, the evolutionary association rule step proposes a multianalysis technique for extending a GA in order to enhance the rule optimization

result. For that, rule ranking and refinement processes are applied as shown in Figure 3. The multianalysis technique starts with the rule-ranking process that is in charge of automatically analyzing and ranking the interesting association rules and the generated rules from supervised learning algorithms regarding their frequency of occurrence and their statistical information. In the proposed rule-ranking process, the rule occurrence frequency is considered as the highest priority to classify rules, followed by the statistical information, such as the fitness function weight. Then, the rule refinement process is performed to derive the set of well-performed rules. This process begins with finding a user who is related to generated rules, loading the user profile of the corresponding user from the ontology-based context model and inferring the well-performed rules that could significantly match with the profile.

**3.2. Rule Transformation Module.** The second module in the proposed approach is the rule transformation module since generating IF-THEN association rules is not enough. This module aims to obtain Jena rules from the well-performed

```

grammar org.xtext.example.mydsl.MyDsl with org.eclipse.xtext.common.Terminals
generate myDsl "http://www.xtext.org/example/mydsl/MyDsl"
import "http://www.eclipse.org/emf/2002/Ecore" as.ecore
RuleList:
    'rules' name = STRING
    '['
    rules += Rule (',' rules += Rule)*
    ']'
;
Rule:
    'rule' name = STRING
    '{'
    concepts += Concepts (',' concepts += Concepts)*
    '}'
;
Concepts:
    Antecedent | Consequent
;

```

FIGURE 6: A fragment of the textual description of the DSL metamodel.

association rules in order to reason over an ontology-based context model and to infer high-level knowledge using the Jena inference engine. The motivation to opt for Jena is its built-in support for rule-based inference over RDF and OWL [28]. In this sense, the rule transformation keeps track of the well-performed association rule set, generated in the previous module, and transforms them into Jena rules according to the rule syntax of Jena. In the following, rules in Jena format induced from the well-performed association rules are loaded by the module into the rule knowledge base for an automatic enrichment purpose at runtime. The rule transformation module targets design time and runtime. At design time, a rule translation is ensured, where a metamodel for Jena rules is proposed to specify the abstract syntax of a Jena rule. Figure 5 shows the metamodel for Jena rules. A rule set consists of rules. A Jena rule has a name and is composed of concepts. A concept is interpreted as an antecedent and as consequent. Therefore, a Jena rule can contain one or more atoms in the antecedent part and only one atom in the consequent part. Each atom has an attribute, type, and value. Then, the proposed metamodel is translated into a domain-specific language (DSL) metamodel. The latter is defined as a textual structure using Xtext. A fragment of the DSL metamodel is depicted in Figure 6.

During runtime, the transformation module performs two paramount phases: (i) instantiation of the DSL metamodel and (ii) generation of the Jena rules on the basis of the DSL model as depicted in Figure 1.

**3.2.1. DSL Rule Instantiation.** The DSL rule instantiation phase is dedicated to introduce a DSL model, which is a formal rule specification of the defined DSL metamodel. The DSL model represents the well-performed obtained association rules. Here, the rule instantiation is automatically performed at runtime.

**3.2.2. Jena Rule Generation.** The Jena rule generation phase is considered for automatically generating the corresponding Jena rules based on the DSL model using Xtend as

a transformation language. Therefore, a set of Jena rules is generated from the DSL model while the semantic of the DSL metamodel is well-defined.

## 4. Experimental Results

In this experimental evaluation, we provide an example of generated rules by our proposed approach. In addition, we discuss the effectiveness of our proposed approach. For this, a range of experiments was carried out to investigate the effectiveness in terms of the number of rules, performance, and computational time.

**4.1. Experimental Setup.** All experiments were conducted on six benchmark datasets of varying complexity. The name, number of instances, number of attributes, and number of class labels in each dataset acquired from the UCI Machine Learning Repository are included in Table 1.

For evaluation purposes, we utilized standard open-source implementations of supervised learning algorithms and association rule mining algorithms provided by Weka [29] in a 10-fold cross-validation evaluation protocol in order to get accurate results for all datasets. In the 10-fold cross-validation protocol, the entire benchmark datasets are partitioned into ten parts of equal size, nine parts of them are used at a time for training, and the remaining one is used for testing. The process is repeated ten times, with different partitions used as training data and test data. In addition, we set the minimum expected weighted confidence threshold to 1 and support threshold to 0.5 in order to generate the results described in the next experiments. Furthermore, the various GA parameters were selected. Crossover and mutation probabilities were taken, respectively, as 0.5 and 0.01. The size of the initial population depends on the number of the generated rules from the benchmark datasets. Thus, the initial population size ranges from 14 to 439 and the maximum number of generations is set to 100.

**4.2. Experimental Metrics.** For analyzing the effectiveness of the proposed approach, some well-known metrics were used. Therefore, the following parameters were considered for these metrics:

- (i) True positive (TP) is the number of rules positively predicted that is actually positive.
- (ii) True negative (TN) is the number of rules negatively predicted that is actually negative.
- (iii) False positive (FP) is the number of rules positively predicted that is actually negative.
- (iv) False negative (FN) is the number of rules negatively predicted that is actually positive.

Based on the previous parameters, the following metrics were proposed for the evaluation:

- (i) Precision reflects the ability of an approach to return relevant rules among a set of irrelevant and relevant rules. The precision can be computed by equation.

- (1) *Input*: DS a preprocessed candidate dataset
- (2) *Output*: Decision trees
- (3) *Begin*
- (4) Apply J48 decision tree to learn on the DS and get the decision tree-based training model
- (5) Apply random tree to learn on the DS and get the random tree-based training model
- (6) Extract decision trees from both training models
- (7) *End*

ALGORITHM 1: Hybrid with decision tree and random tree.

TABLE 1: Characteristics of datasets used in experiments.

Dataset	Number of instances	Number of attributes	Number of classes
Breast	286	10	2
Heart	270	13	4
Hepatitis	155	19	2
Weather	14	5	3
Iris	150	4	3
Adult	48842	15	2

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2)$$

- (ii) Recall reflects the ability of an approach to return relevant rules only. The recall is defined as given in equation.

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3)$$

- (iii) Accuracy reflects the ability of an approach to return the accurate rules over the all rules made in the dataset as can be seen in equation.

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \quad (4)$$

Apart from these metrics, we considered the computational time that reflects the average time required for generating the set of IF-THEN association rules in terms of second.

**4.3. Rules for Knowledge Base Enrichment.** We present an example of context-aware rules generated from the weather dataset mentioned above using our approach. The candidate dataset contains four condition attributes, such as outlook, temperature, humidity, windy, and one decision attribute, that is, “play.” It is considered to generate rules that could help in making decisions regarding whether a user could go outside for playing or not. First, we show, in Table 2, a sample of generated IF-THEN rules obtained through training both decision tree and random tree algorithms. Looking more closely at Table 2, we can notice that the antecedent part of the rules reflects users’ contextual knowledge, and the consequent part represents their associated behavioral actions. Hence, real-time users’ environmental contexts such as outlook, wind, humidity, or temperature are used to make the play decision.

Then, we illustrate, in Tables 3 and 4, a sample of well-performed rules extracted from the generated rules through the optimization process considered in the proposed approach. As illustrated in Tables 3 and 4, the well-performed discovered association rules vary from one user to another according to their current profile. For example, a user with a pollen allergy, on a windy day, cannot go outside since the wind can trigger pollen allergy symptoms most. As a result, association rules, such as R9 and R10, which stated a day is windy (windy = true) and a user could have fun playing outside (play = yes), are overlooked for the user with a pollen allergy.

Finally, after the automatic DSL metamodel instantiation and transformation, we list, in Tables 5 and 6, a sample of the well-performed rules transformed to Jena rules to enrich the rule knowledge base and improve the decision-making process at runtime. Therefore, we can conclude that our proposed approach is capable of generating rules that could enrich a rule knowledge base.

#### 4.4. Effectiveness Analysis

**4.4.1. Rule Analysis.** In the first experiment, the number of association rules generated by the proposed approach and the traditional algorithms of association rule mining, such as apriori and FP growth, was evaluated. Figures 7–12 illustrate the number of generated rules for the different candidate datasets.

Observing Figures 7–12, we can see that the traditional algorithms generate the highest number of rules, while our approach generates the lowest number of rules on all candidate datasets. In particular, our approach generates 4,103 association rules on average against 6 datasets, whereas the apriori and FP growth algorithms derive 9,211 and 10,582 on average, respectively. Thus, the results show that the number of generated association rules using traditional algorithms is large and huge. The reason beyond these results is that traditional algorithms simply take into account all



TABLE 2: A sample of generated rules from the weather dataset.

Rules	IF-THEN rules
R1	IF outlook = overcast THEN play = yes
R2	IF outlook = rainy AND windy = false THEN play = yes
R3	IF outlook = rainy AND windy = true THEN play = no
R4	IF outlook = sunny AND humidity = high THEN play = no
R5	IF outlook = sunny AND humidity = normal THEN play = yes
R6	IF humidity = normal AND windy = false THEN play = yes
R7	IF humidity = high AND outlook = rainy AND windy = false THEN play = yes
R8	IF humidity = high AND outlook = rainy AND windy = true THEN play = no
R9	IF humidity = normal AND windy = true AND temperature = hot THEN play = yes
R10	IF humidity = normal AND windy = true AND temperature = cool AND outlook = overcast THEN play = yes

TABLE 3: A sample of well-performed rules for a user with no allergy history.

Rules	IF-THEN rule	Occurrence frequency	Fitness value
R1.1	IF outlook = overcast THEN play = yes	1	0.743
R1.2	IF humidity = normal AND windy = false THEN play = yes	1	0.686
R1.3	IF humidity = normal AND windy = true AND temperature = cool AND outlook = overcast THEN play = yes	1	0.675
R1.4	IF humidity = normal AND windy = true AND temperature = hot THEN play = yes	1	0.611

TABLE 4: A sample of well-performed rules for a user with a pollen allergy.

Rules	IF-THEN rule	Occurrence frequency	Fitness value
R2.1	IF outlook = overcast THEN play = yes	1	0.743
R2.2	IF humidity = normal AND windy = false THEN play = yes	1	0.686
R2.3	IF outlook = sunny AND humidity = high THEN play = no	2	0.611
R2.4	IF outlook = rainy AND windy = true THEN play = no	1	0.611

TABLE 5: A sample of well-performed rules in Jena format for a user with no allergy history.

Rules	Jena rules
R1.1	(rule R1: (?outlookValue uni:outlook "overcast") → (?playValue uni:play "yes"))
R1.2	(rule R2: (?humidityValue uni:humidity "normal") (?windyValue uni:windy "false") → (?playValue uni:play "yes"))
R1.3	(rule R3: (?humidityValue uni:humidity "normal") (?windyValue uni:windy "true") (?temperatureValue uni:temperature "cool") (?outlookValue uni:outlook "overcast") → (?playValue uni:play "yes"))
R1.4	(rule R4: (?humidityValue uni:humidity "normal") (?windyValue uni:windy "true") (?temperatureValue uni:temperature "hot") → (?playValue uni:play "yes"))

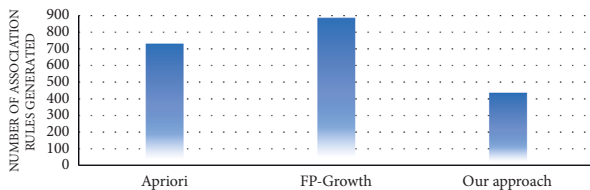


FIGURE 7: Rule analysis for the breast dataset.

combinations of attributes while generating rules. In contrast, the results indicate that the number of generated rules by our approach is small. The downtrend reveals that our approach could generate the minimum number of association rules comparing the apriori and FP growth and could keep the number of discovered rules as small as possible. As a result, for a high confidence value, traditional algorithms satisfy significantly more associations than our proposed

approach. Therefore, our proposed approach could generate a reasonably smaller number of association rules compared with traditional association rule mining algorithms since it abandons the redundant rules and retains the nonredundant rules.

**4.4.2. Performance Analysis.** In the second experiment, we discussed the effectiveness of our approach in terms of performance measures such as precision, recall, and accuracy. For this purpose, we compare the performance of the proposed approach with well-known supervised learning algorithms, namely, NB, JRip, and decision table (DT). The reason for selecting these algorithms is that they generate rule-based classifiers and have high performance compared with other algorithms [30]. We also compare the performance of our approach with some similar state-of-the-art works, including Hong et al. [14], Sarker [16], and Basha

TABLE 6: A sample of well-performed rules in Jena format for a user with a pollen allergy.

Rules	Jena rules
R2.1	(rule R1: (?outlookValue uni:outlook "overcast") → (?playValue uni:play "yes"))
R2.2	(rule R2: (?humidityValue uni:humidity "normal") (?windyValue uni:windy "false") → (?playValue uni:play "yes"))
R2.3	(rule R3: (?outlookValue uni:outlook "sunny") (?humidityValue uni:humidity "high") → (?playValue uni:play "no"))
R2.4	(rule R4: (?outlookValue uni:outlook "rainy") (?windyValue uni:windy "true") → (?playValue uni:play "no"))

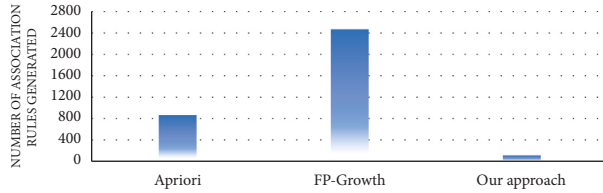


FIGURE 8: Rule analysis for the heart dataset.

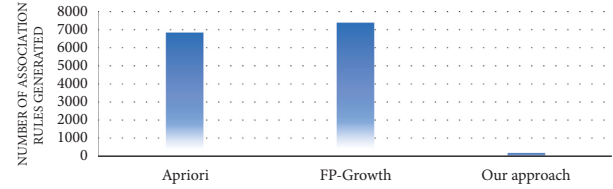


FIGURE 9: Rule analysis for the hepatitis dataset.

[17]. Experimental results on performance measures are highlighted in Figures 13–18.

Our observations on candidate datasets show that our approach consistently outperforms the compared supervised learning algorithms for generating association rules by maximizing the precision and recall. In addition, our observations reveal that our proposed approach achieved better accuracy than NB, JRip, and DT on heart, hepatitis, weather, and iris datasets. For instance, the accuracies for NB, JRip, and DT on the iris dataset are 95.33%, 96%, 95.3%, and 92.66%, respectively, whereas for the proposed approach the accuracy is 98%. Moreover, the obtained results confirm that our approach has outstanding performance compared with state-of-the-art works. For instance, we achieve an approximately 2% and 3.2% accuracy gain compared with the works of Hong et al. [14], Sarker [16], and Basha [17] when dealing with breast and adult datasets, respectively. Thus, obtained results proved that our proposed approach tends to get reasonably high accuracy on all datasets. Therefore, we can conclude that our proposed approach is more effective relative to the compared supervised learning algorithms and state-of-the-art works while generating association rules since we capture association rules from both more performant supervised learning algorithms that lead to improve the performance results.

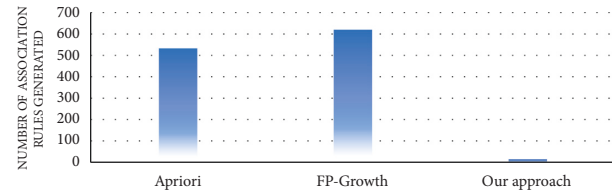


FIGURE 10: Rule analysis for the weather dataset.

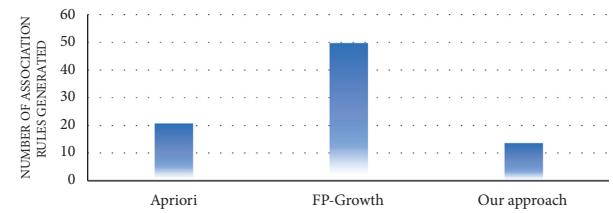


FIGURE 11: Rule analysis for the iris dataset.

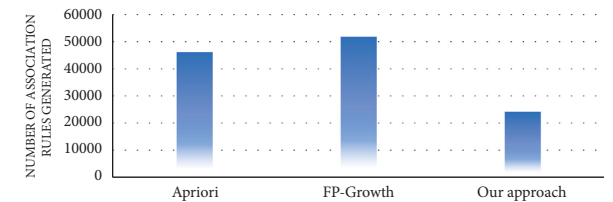


FIGURE 12: Rule analysis for the adult dataset.

4.4.3. *Computational Time Analysis.* In the third experiment, we compared and analyzed the computational time of the proposed approach, the supervised learning algorithms mentioned earlier, and the state-of-the-art works [14, 16]. In this experiment, we did not consider the work presented by Basha [17] since its GA is not publicly available. To this end, the sizes of the datasets were fixed since the computational time may vary based on the dataset size. Table 7 illustrates the time consumed by the proposed approach for generating association rules against the selected supervised learning algorithms and the state-of-the-art works in all datasets.

From the illustrated results, the average time spent on each dataset to generate a set of association rules is 0.38 seconds in the proposed approach. The experimental results

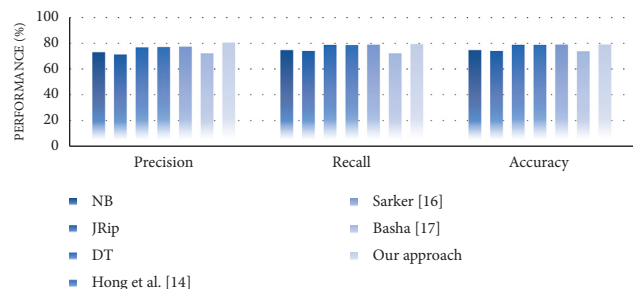


FIGURE 13: Performance analysis for the breast dataset.

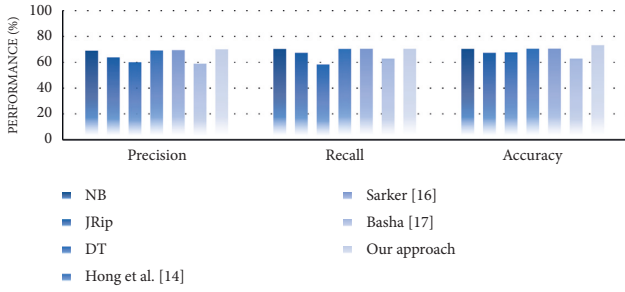


FIGURE 14: Performance analysis for the heart dataset.

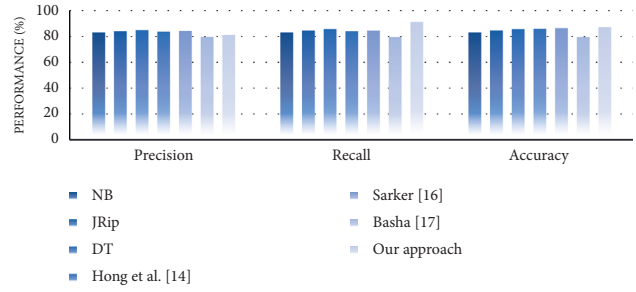


FIGURE 18: Performance analysis for the adult dataset.

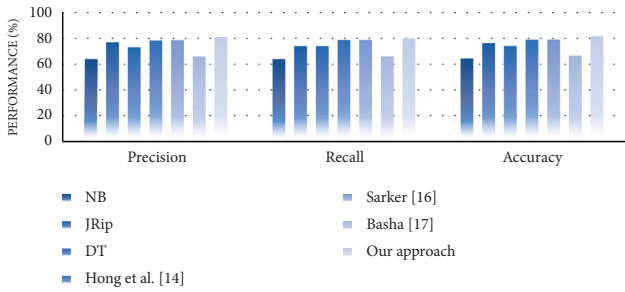


FIGURE 15: Performance analysis for the hepatitis dataset.

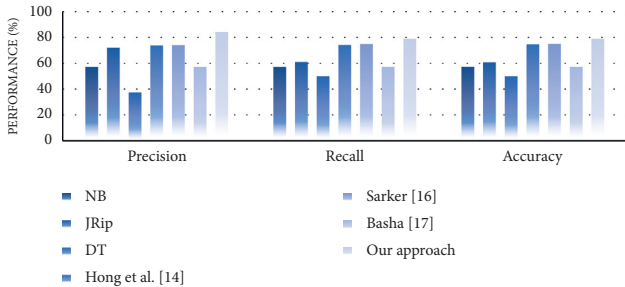


FIGURE 16: Performance analysis for the weather dataset.

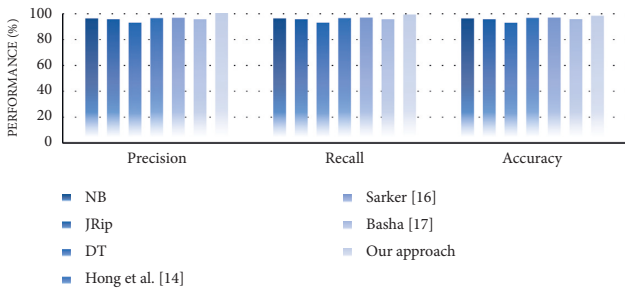


FIGURE 17: Performance analysis for the iris dataset.

clearly show that the computational time of the proposed approach is slightly increased compared with the selected supervised learning algorithms and with the state-of-the-art works, but the computational time on the whole is not much different. The slight increase can be explained by the fact that we apply the GA for optimization on the basis of both

TABLE 7: Computational time analysis for our six datasets in seconds.

Dataset	NB	JRip	DT	Hong et al. [14]	Sarker [16]	Our approach
Breast	0.05	0.04	0.02	0.03	0.05	0.06
Heart	0.02	0.05	0.06	0.04	0.06	0.10
Hepatitis	0.05	0.04	0.03	0.03	0.05	0.08
Weather	0.02	0.02	0.05	0.04	0.06	0.16
Iris	0.02	0.01	0.02	0.02	0.04	0.06
Adult	1.13	1.48	1.94	1.08	1.10	1.82
Average	0.29	0.27	0.35	0.21	0.23	0.38

supervised learning algorithms, which could slightly increase the time complexity of the rule generation process compared with the selected supervised learning algorithms.

### 5. Discussion

We discuss the encouraging obtained results of the proposed approach. First, our approach outperformed compared to the well-known traditional association rule mining algorithms in terms of the number of rules by eliminating the redundant generation for each dataset. It generated, on average far, fewer association rules than those generated by the traditional algorithms included in the comparison. Thus, it provided a reasonably smaller number of rules on smaller and bigger datasets compared to the traditional algorithms. This is due to the fact that our approach is based on hybrid supervised learning that takes advantage of machine learning to extract the relevant relationships and to eliminate the redundancy while generating association rules. In brief, our approach significantly reduces the total number of generated rules and outputs a well-performed set of association rules using the GA. Such nonredundant and well-performed association rule generation makes our approach more effective and can be used to automatically improve a decision-making process regarding changes occurring in users' surrounding environments at runtime. Second, our approach exceeded the well-known selected supervised learning algorithms and certain state-of-the-art works in terms of precision, recall, and accuracy. In particular, our approach achieved the best accuracy on heart, hepatitis, weather, and iris datasets among all selected supervised learning algorithms. However, it got slightly worse results in terms of accuracy on breast and adult datasets since

imbalanced datasets may lead to slightly worse accuracy results. Even though not achieving the best accuracy results on these datasets, our approach achieves almost reasonably high accuracy in most datasets. Moreover, we must stress that the main advantage of our approach in contrast to the compared state-of-art works is its ability to achieve a good trade-off between precision, recall, and accuracy on multiple datasets. Third, our approach could be somewhat more time-consuming than the selected supervised learning algorithms and the compared state-of-the-art works. The slight increase in time is reasonable due to the hybridization and optimization. However, this loss of the little running efficiency will result in a significant improvement in the quality of the results in terms of the number of rules, precision, recall, and accuracy. Regarding the time-consuming results, it is noted that the effectiveness of the proposed approach does have room for improvement, which is also the need for further improvement in the future work.

Overall, the findings of the experimental study reveal that our approach can (i) effectively minimize the issues of redundant rule generation, (ii) provide a promising performance and a high accuracy to extract a concise set of association rules, and (iii) take slightly more time for generating a set of well-performed rules.

## 6. Conclusions

This work aimed to propose an approach for the decision-making improvement to face dynamic environments at runtime, where the enrichment of a rule knowledge base is performed through automatically generating and transforming context-aware rules. The proposed approach focused first on generating nonredundant association rules with a hybrid supervised learning mechanism, second on optimizing association rules using an extended GA, and finally on transforming association rules to context-aware rules in Jena syntax. As part of this work, we presented an experimental evaluation to assess the effectiveness of the proposed approach. The results showed that the proposed approach has the potential for achieving the best results in terms of the number of rules, precision, recall, and accuracy among all compared algorithms. Moreover, these results pinpointed the limitation of the proposed approach in terms of time-consuming that is mainly due to hybridization and optimization. Furthermore, in the near future, we intend to apply our approach to real scenarios for a context-aware system.

## Data Availability

The public datasets were collected from the UCI Machine Learning Repository. Requests for dataset access should be made to <https://archive.ics.uci.edu/ml/index.php/>.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References

- [1] R. C. A. Da Rocha and M. Endler, *Context Management for Distributed and Dynamic Context-Aware Computing*, Springer Science & Business Media, Berlin/Heidelberg, Germany, 2012.
- [2] C. K. Chang, "Situation analytics: a foundation for a new software engineering paradigm," *Computer*, vol. 49, no. 1, pp. 24–33, 2016.
- [3] T. Zhao, "The generation and evolution of adaptation rules in requirements driven self-adaptive systems," in *Proceedings of the 2016 IEEE 24th International Requirements Engineering Conference (RE)*, pp. 456–461, IEEE, Beijing, China, 2016, September.
- [4] Y. Liu, W. Zhang, and W. Jiao, "A generative genetic algorithm for evolving adaptation rules of software systems," in *Proceedings of the 8th Asia-Pacific Symposium on Internetware*, pp. 103–107, New York, New York, United States, 2016, September.
- [5] I. H. Sarker, "Context-aware rule learning from smartphone data: survey, challenges and future directions," *Journal of Big Data*, vol. 6, no. 1, pp. 95–25, 2019.
- [6] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, 1989.
- [7] J. J. Carroll, I. Dickinson, C. Dollin, D. Reynolds, A. Seaborne, and K. Wilkinson, "Jena: implementing the semantic web recommendations," in *Proceedings of the 13th international World Wide Web conference on Alternate track papers & posters*, pp. 74–83, New York, New York, United States, 2004, May.
- [8] A. Asuncion and D. Newman, *UCI Machine Learning Repository*, 2007.
- [9] M. Gabroveanu and I. M. Diaconescu, "Extracting semantic annotations from Moodle data," *Proceedings of the 2nd East European Workshop on Rule-Based Applications (RuleApps 2008) at the 18th European Conference on Artificial Intelligence (ECAI 2008)*, pp. 1–5, New York, New York, United States, 2008, July.
- [10] J. Kaliappan and S. M. Sai, *Weblog and Retail Industries Analysis Using a Robust Modified Apriori Algorithm*, 2019.
- [11] K. Davagdorj and K. H. Ryu, "Association rule mining on Head and Neck Squamous Cell Carcinoma Cancer using FP-Growth algorithm," *Proceedings of the International Conference on Information, System and Convergence Applications*, 2018.
- [12] S. Asadianfam, H. Kolivand, and S. Asadianfam, "A new approach for web usage mining using case based reasoning," *SN Applied Sciences*, vol. 2, no. 7, pp. 1–11, 2020.
- [13] N. H. Miswan, I. M. Sulaiman, C. S. Chan, and C. Guan Ng, "Association rules mining for Hospital readmission: a case study," *Mathematics*, vol. 9, no. 21, p. 2706, 2021.
- [14] J. Hong, E. H. Suh, J. Kim, and S. Kim, "Context-aware system for proactive personalized service based on context history," *Expert Systems with Applications*, vol. 36, no. 4, pp. 7448–7457, 2009.
- [15] S. Zulkernain, P. Madiraju, S. I. Ahamed, and K. Stamm, "A mobile intelligent interruption management system," *Journal of Universal Computer Science*, 2010.
- [16] I. H. Sarker, "A machine learning based robust prediction model for real-life mobile phone data," *Internet of Things*, vol. 5, pp. 180–193, 2019.
- [17] M. S. Basha, "Early prediction of Cardio Vascular Disease by performing associative classification on medical datasets and using genetic algorithm," in *Intelligent Computing and*

- Innovation on Data Science*, pp. 393–402, Springer, Singapore, 2021.
- [18] A. Mahmood, K. Shi, S. Khatoon, and M. Xiao, “Data mining techniques for wireless sensor networks: a survey,” *International Journal of Distributed Sensor Networks*, vol. 9, no. 7, p. 406316, 2013.
  - [19] A. A. Freitas, “Understanding the crucial differences between classification and discovery of association rules: a position paper,” *AcM SIGKDD Explorations Newsletter*, vol. 2, no. 1, pp. 65–69, 2000.
  - [20] J. R. Quinlan, H. Saad, and N. Nagarur, “Induction of decision trees,” *Machine Learning*, vol. 1, no. 1, pp. 81–106, 1986.
  - [21] I. H. Witten and E. Frank, “Data mining: practical machine learning tools and techniques with Java implementations,” *Acm Sigmod Record*, vol. 31, no. 1, pp. 76–77, 2002.
  - [22] H. Saad and N. Nagarur, *Data Mining Techniques in Predicting Breast Cancer*, arXiv preprint arXiv:2011.11088, 2020.
  - [23] A. Suresh, R. Udendhran, and M. Balamurgan, “Hybridized neural network and decision tree based classifier for prognostic decision making in breast cancers,” *Soft Computing*, vol. 24, no. 11, pp. 7947–7953, 2020.
  - [24] J. J. Grefenstette, “Genetic algorithms and machine learning,” *Proceedings of the sixth annual conference on Computational learning theory*, pp. 3–4, 1993.
  - [25] R. Haldulakar and J. Agrawal, “Optimization of association rule mining through genetic algorithm,” *International Journal on Computer Science and Engineering*, vol. 3, no. 3, pp. 1252–1259, 2011.
  - [26] K. N. V. D. Sarath and V. Ravi, “Association rule mining using binary particle swarm optimization,” *Engineering Applications of Artificial Intelligence*, vol. 26, no. 8, pp. 1832–1840, 2013.
  - [27] H. R. Qodmanan, M. Nasiri, and B. Minaei-Bidgoli, “Multi objective association rule mining with genetic algorithm without specifying minimum support and minimum confidence,” *Expert Systems with Applications*, vol. 38, no. 1, pp. 288–298, 2011.
  - [28] S. Hussain and S. S. R. Abidi, “Ontology driven CPG authoring and execution via a Semantic Web framework,” in *2007 40th Annual Hawaii International Conference on System Sciences (HICSS’07)*, p. 135, IEEE, 2007.
  - [29] Weka, “Weka 3: machine learning software in Java (n.d.),” 2021, <https://www.cs.waikato.ac.nz/ml/weka/>.
  - [30] L. J. Rubini and P. Eswaran, “Generating comparative analysis of early stage prediction of Chronic Kidney Disease,” *International Journal of Modern Engineering Research (IJMER)*, vol. 5, no. 7, pp. 49–55, 2015.



## Research Article

# An Efficient Data Classification Decision Based on Multimodel Deep Learning

Wenjin Hu <sup>1,2</sup>, Feng Liu <sup>1,2</sup> and Jiebo Peng <sup>1,2</sup>

<sup>1</sup>School of Computer Science, Xi'an Polytechnic University, Xi'an, Shaanxi, China

<sup>2</sup>Shaanxi Key Laboratory of Clothing Intelligence and Computation, Xi'an, Shaanxi, China

Correspondence should be addressed to Feng Liu; liufeng@xpu.edu.cn

Received 28 December 2021; Accepted 20 April 2022; Published 4 May 2022

Academic Editor: Daqing Gong

Copyright © 2022 Wenjin Hu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A single model is often used to classify text data, but the generalization effect of a single model on text data sets is poor. To improve the model classification accuracy, a method is proposed that is based on a deep neural network (DNN), recurrent neural network (RNN), and convolutional neural network (CNN) and integrates multiple models trained by a deep learning network architecture to obtain a strong text classifier. Additionally, to increase the flexibility and accuracy of the model, various optimizer algorithms are used to train data sets. Moreover, to reduce the interference in the classification results caused by stop words in the text data, data preprocessing and text feature vector representation are used before training the model to improve its classification accuracy. The final experimental results show that the proposed model fusion method can achieve not only improved classification accuracy but also good classification effects on a variety of data sets.

## 1. Introduction

Text classification originated in the 1950s. From the initial expert-based system to traditional machine learning approaches and now deep learning methods, text classification technology is gradually maturing [1]. In recent years, deep learning has been a hot topic in academic research. Great breakthroughs have been made in various fields by using deep learning technology, such as natural language processing, computer vision, and reinforcement learning. Text classification based on deep learning is both a trending subject and a long-term challenge for researchers.

In a study on text classification, Rocchio [2] first proposed the Rocchio text classification algorithm, which uses a training set to construct a prototype vector for each class and allocates an input document to a certain class by calculating the similarity between all documents in the training set and the prototype vector. This method is easy to implement and calculate. However, it does not perform well in tasks with multiple categories and is mostly suitable for document classification problems with fewer categories. Based on the Rocchio algorithm, Somya and Srinivasa [3] proposed a

Rocchio algorithm with a hierarchical structure. This algorithm adopts the term frequency-inverse document frequency (TF-IDF) feature extraction method to conduct large-scale multilabel text classification for hierarchical data sets and has a good classification effect on such data sets. However, it has poor performance on multitype data sets. Schapire [4] first proposed the boosting classification algorithm based on the idea of model fusion. This algorithm mainly fuses multiple weak learners into a strong learner, achieving an improved classification effect; then, he proposed the bagging algorithm [5]. Among these weak learners, random forests, boosting trees, and gradient boosting decision trees (GBDTs) [6] are the basic models used when combining decision trees with boosting and bagging algorithms; this approach provides a significantly improved classification effect. Based on the idea of boosting, Bloehdorn and Hotho [7] proposed an adaptive boosting algorithm with semantics. The model algorithm uses an integrated learning method to improve the stability and accuracy of text classification. However, the number of trained models is large; the calculations are very complicated; and the interpretability between models is reduced.

Kowsari [8] and others proposed a random multimodel classification method, which uses the network learning architectures of a deep neural network (DNN), recurrent neural network (RNN), and convolutional neural network (CNN) to randomly generate numbers of hidden layers and neurons for each model and obtain predictions through majority voting. As a result, this method has achieved improved text classification accuracy. However, since the numbers of hidden layers and neurons are randomly generated, the network structure generated each time is different, making the model training process difficult and the calculations very complicated. References [9–11] proposed a widely used Bayes classifier, which is a single classifier. It has a good classification effect on text data, and the calculations are fast and easy to implement. However, the Bayesian classifier performs poorly on text with sparse data. Therefore, Kim et al. [12] developed a method and a strong data distribution assumption to alleviate the problem regarding poor prediction for sparse data and solve the poor classification problem for sparse text data. Another powerful classifier is the support vector machine (SVM) [13], which uses a flexible and diverse kernel method to project data into a high-dimensional space, thereby using the hyperplane in the high-dimensional space to classify the data; the classification effect of this method is relatively good, but the effect of a single classifier is reduced when used on multiple data sets.

The traditional single classifier, which can be used to achieve good classification results, is used to train the network model for a specific data set. While the single classifier cannot handle various data sets, the integrated model can solve the problem of accuracy under various data sets, but the training of the integrated model is complex and time-consuming. The traditional single classification model and integrated model are improved in this paper, and the data sets are trained by combining multiple network models. First, a variety of optimizer algorithms are used to address text feature problems such as long texts and sparse texts. Second, the learning rate in the training process of the model is adjusted by the optimizer so that the trained model can handle data sets with different text characteristics and increase the flexibility of the model.

Improvements are made on the basis of three network architectures: DNN, RNN, and CNN. In the DNN network architecture, the BP algorithm and many optimizer algorithms are combined to train the model. In the RNN network architecture, the GRU [14] network with horizontal and vertical learning, multiple optimizer algorithms, and the ReLU [15] activation function training model are designed. In the CNN network architecture, multilayer convolution and pooling are used to extract text features and combined with a variety of optimizer algorithms [16] to train the model. Finally, the improved network architectures are used to obtain a powerful text classification model by using a fusion strategy and parallel training method.

## 2. Materials and Methods

The text classification process in this paper includes the following steps: text data preprocessing, text vector

representation, text feature extraction, optimizer selection, model design, model fusion, and model evaluation. Figure 1 shows the flowchart of the proposed method from preprocessing to model evaluation. The following describes the detailed process of the whole text classification in this article.

*2.1. Text Preprocessing.* The main purpose of text preprocessing is to clean the original text data. Most text and document data sets contain many unnecessary words, such as stop words, spelling errors, and slang. When determining word frequency statistics, these noisy data and unnecessary features adversely affect the performance of the models and the classification results. In this paper, the stop vocabulary list in the NLTK corpus is incorporated for word removal and processing. Additionally, regular expressions are used to remove spaces and some special characters to reduce the calculation cost and classification error of the model training.

*2.2. Word Representation.* Word representation converts a text string into a numerical vector that can be processed by a computer. The commonly used text vector representation methods are the bag-of-words [17] method, word2vec [18], and the GloVe [19] method. This paper uses the GloVe model for text vector representation. The basic idea is to construct a co-occurrence matrix through a corpus and then learn word vectors based on the co-occurrence matrix and the GloVe model. The model uses matrix decomposition with latent semantic analysis (LSA) derived global features and word2vec local context information to transform the matrix. Together, this approach not only realizes the global statistical features in the corpus but also uses local context features to represent vectors. When using the GloVe model for training, the word vector for the given text is calculated by introducing a loss function such as

$$\mathcal{J} = \sum_{i,j}^N f(X_{i,j}) (v_i^T v_j + b_i + b_j - \log(X_{i,j}))^2, \quad (1)$$

where  $v_i$  and  $v_j$  are the word vectors of word  $i$  and word  $j$ , respectively;  $b_i$  and  $b_j$  are bias terms;  $f$  is a weight function;  $\log(X_{i,j})$  represents the number of times word  $j$  appears in the context of word  $i$ ;  $N$  is the size of the vocabulary; and  $N * N$  is the dimensionality of the co-occurrence matrix. Additionally, the weight function of  $f$  is defined; this is done to prevent some weights from being too large. Additionally, some weights are too small, which affects the text classification accuracy, so function  $f(x)$  is defined as follows:

$$f(x) = \begin{cases} \left(\frac{x}{x_{\max}}\right)^\alpha, & x < x_{\max}, \\ 1, & \text{Other.} \end{cases} \quad (2)$$

*2.3. Feature Extraction.* The commonly used feature extraction methods are TF-IDF [20] and  $N$ -gram [21]. When

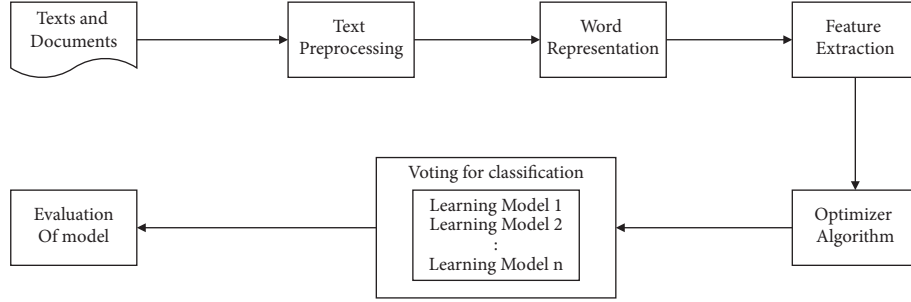


FIGURE 1: Overview of the text classification pipeline.

TF-IDF is used to extract text features, the sequence relationship between words cannot be captured, so the features of the text cannot be fully extracted. Therefore, this paper uses the  $N$ -gram method to extract text features. In text data, a sequence or a sentence is composed of  $m$  words. Then, the probability of the occurrence of the  $m$ -th word is related to the previous  $m-1$  words, and the probability value  $p(w_1, w_2, \dots, w_m)$  is calculated. According to the chain rule, the final calculation result is shown in the following equation:

$$p(w_1, w_2, \dots, w_m) = p(w_1) * p(w_2|w_1) * p(w_3|w_1, w_2) \dots p(w_m|w_1, \dots, w_{m-1}) \quad (3)$$

Compared with a 1-gram and a 2-gram, the  $N$ -gram method provides extracted text features that can detect more information.

#### 2.4. Optimizer Algorithm

**2.4.1. An Optimization Algorithm for Vibration Reduction Based on RMSProp.** When training the parameters of the model, the choice of the learning rate affects the efficiency and performance of the model. If the learning rate is too large, violent oscillations will occur when calculating the gradient, resulting in failure to converge to the global optimal solution, and if the learning rate is too small, the training speed will become very slow. The current method is the simulated annealing algorithm. When training model parameters, a threshold is defined for the learning rate change range, and the learning rate is adjusted. However, this approach requires a threshold to be defined in advance, and it cannot adapt to changing text data types. To solve the problem regarding learning rate threshold changes and oscillations, this paper uses the RMSProp algorithm based on Nesterov momentum. The algorithm first initializes the learning rate and continuously dynamically updates the learning rate during the training process to prevent this rate from being excessive or small during the training process, as this would affect the training results. The specific main steps are as follows. First, a minibatch containing  $m$  samples from the training set is formed, and the gradient sum and average of these small samples are calculated, as in equation (4). At the same time, an exponential decay coefficient such as equation (5) is used to control the amount of historical information, that is, the cumulative gradient, to dynamically

update the size of the learning rate (equation (6)) and the model parameters (equation (7)).

$$g = \frac{1}{m} \nabla_{\theta} \sum_i \mathcal{L}(f(x^{(i)}; \bar{\theta}), y^{(i)}), \quad (4)$$

$$r = \rho r + (1 - \rho) g * g, \quad (5)$$

$$v = \alpha v - \frac{\epsilon}{\sqrt{r}} * g, \quad (6)$$

$$\theta = \theta + v, \quad (7)$$

$$\bar{\theta} = \theta + \alpha v, \quad (8)$$

where  $r$  represents the exponential decay coefficient,  $v$  represents the update of the calculation decay,  $\theta$  represents the update of the parameters,  $g$  represents the gradient, and  $\alpha$  represents the momentum coefficient. At the same time, to reduce the violent oscillations observed when training the model parameters, an idea based on Nesterov momentum is adopted, and a momentum coefficient is added before calculating the gradient to slow the oscillations, as shown in equation (8). Therefore, the optimizer of the RMSProp algorithm combined with Nesterov momentum improves the training speed and accuracy of the model.

**2.4.2. An Optimal Deviation Correction Algorithm Based on Adam.** Adam is a simple and computationally efficient optimization algorithm that can overcome the problems encountered in large data sets and high-dimensional parameter spaces. Among traditional classification methods, such as naive Bayes classifiers and SVMs, for text data with large data sets, it is necessary to consider the memory consumption of the naive Bayes model training process and the model oscillation problem, while SVMs do not possess this problem. A kernel method puts text data into a high-dimensional space, and the resulting model has many parameters and is highly complicated to calculate. Therefore, the Adam algorithm with deviation correction is introduced to solve the shortcomings of traditional methods. The basic idea is that when Adam calculates the gradient, it introduces the first-order moment estimation (equation (9)) and the second-order moment estimation (equation (10)) and then corrects the deviation of the first-order moment (equation



(11) and that of the second-order moment (equation (12)) to address gradient sparseness and unevenness. Equation (9) updates the first-order moment by calculating the gradient  $g$  and the exponential decay rate  $\rho_1$  and simultaneously corrects the deviation through equation (11), speeding up the convergence of the model. Equation (10) introduces the second-order moment, and equation (12) corrects the deviation to improve the model's ability to deal with non-stationary targets. At the same time, the  $\Delta\theta$  parameter (equation (13)) is used to update the value of parameter  $\theta$  in equation (14). This paper uses the aforementioned algorithm in text classification to train the model parameters to reduce memory consumption and simultaneously solve the problem of convex model convergence.

$$s = \rho_1 s + (1 - \rho_1)g, \quad (9)$$

$$r = \rho_2 r + (1 - \rho_2)g * g, \quad (10)$$

$$\hat{s} = \frac{s}{1 - \rho_1^t}, \quad (11)$$

$$\hat{r} = \frac{r}{1 - \rho_2^t}, \quad (12)$$

$$\Delta\theta = -\varepsilon \frac{\hat{s}}{\delta + \sqrt{\hat{r}}}, \quad (13)$$

$$\theta = \theta + \Delta\theta, \quad (14)$$

where  $s$  represents the first-order moment,  $g$  represents the value of the sample gradient,  $\rho_1$  and  $\rho_2$  represent the estimated exponential decay rate,  $\theta$  represents the parameter calculated by the model,  $\delta$  represents a small constant used for numerical stability,  $\varepsilon$  represents the step size, and  $\Delta\theta$  is used to update the parameter  $\theta$ .

#### 2.4.3. An Improved Optimization Algorithm Based on SGD.

Before the introduction of the SGD algorithm, the most commonly used gradient algorithm was the batch gradient descent (BGD) algorithm, which is aimed at an entire data set and calculates the gradient direction for all samples. Although this method can obtain the global optimal solution, when the amount of data is large, the number of required calculations is large, and the calculation speed is slow. To overcome the shortcomings of the BGD method, this paper uses the SGD algorithm. SGD is a widely used optimization algorithm. It is an improvement over the classic gradient descent algorithm. The basic idea is that all training data can be obtained from the training data in each iteration. A random sample is taken to estimate the gradient of the objective function, so the time complexity of the algorithm is greatly reduced, and this approach is applied to large-scale text data sets. When using the SGD algorithm in this paper, a set of text training data is input into the model. The objective function is calculated as follows:

$$\min_{w \in \mathbb{R}^n} F(w) = \frac{1}{m} \sum_i \mathcal{L}(f(x^{(i)}; w), y^{(i)}), \quad (15)$$

where  $\mathcal{L}$  represents the experience loss of the model  $f(x^{(i)}; w)$ ; the model parameter values are calculated as follows:

$$w = \frac{1}{m} \sum_i w_i - \eta_i \nabla \mathcal{L}(f(x^{(i)}; w_i), y^{(i)}). \quad (16)$$

### 2.5. Model Structure and Fusion

**2.5.1. DNN Architecture.** The DNN structure in this paper is designed with an input layer, a hidden layer, and an output layer, as shown in Figure 2. The input layer is the processed text feature vector, and the ReLU activation function is used in the hidden layer (as in equation (17)) to reduce the required number of calculations when using the back-propagation algorithm to update the parameters. At the same time, the dropout algorithm is introduced to solve the problem of gradient disappearance in the training process, and finally, the softmax function, such as equation (18), is used to solve the multiclassification problem when outputting.

$$f(x) = \max(0, x), \quad (17)$$

$$\text{Softmax}(z_i) = \frac{e^{z_i}}{\sum_{c=1}^C e^{z_c}}, \quad (18)$$

where  $z_i$  represents the output value of the  $i$ -th node and  $c$  is the number of output nodes.

**2.5.2. RNN Architecture.** To solve the information loss problem in the traditional RNN network propagation process, the gate structures of LSTM [22] and GRU networks are generally used to retain important information. Because the network parameters of GRU are less than those of LSTM, the gradient disappearance problem can be prevented and reduce the computational complexity and the overfitting of the training data. The GRU network design unit is shown in Figure 3. This method uses a gating mechanism, which contains an update gate and a reset gate. The calculations are shown in equations (19) and (20), and the output vector is calculated via equations (21) and (22). The final RNN architecture is shown in Figure 4. Each GRU unit in the network can learn not only horizontally but also longitudinally to reduce the information loss problem in the communication process.

$$z_t = \sigma(\mathcal{W}_z * [h_{t-1}, x_t]), \quad (19)$$

$$r_t = \sigma(\mathcal{W}_r * [h_{t-1}, x_t]), \quad (20)$$

$$\tilde{h}_t = \tanh(\mathcal{W}_h * [h_{t-1} * r_t, x_t]), \quad (21)$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t, \quad (22)$$

where  $z_t$  represents the update gate vector at time  $t$ ,  $x_t$  represents the input text feature vector,  $\mathcal{W}$  represents the parameter matrix,  $\sigma$  represents the ReLU activation

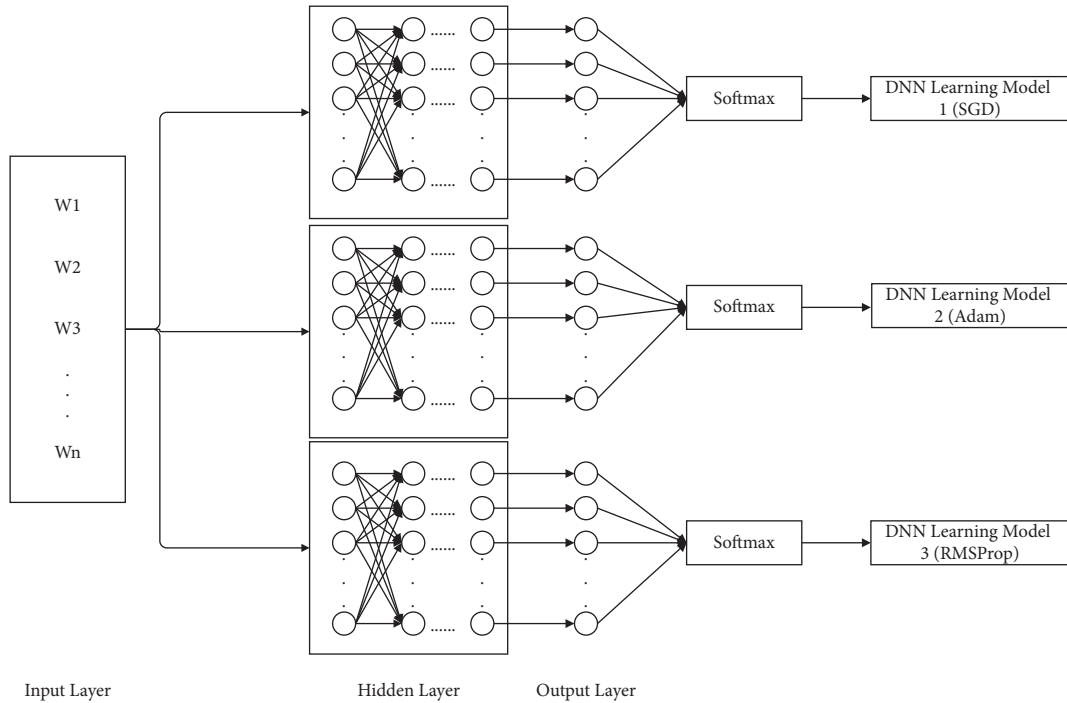


FIGURE 2: DNN architecture.

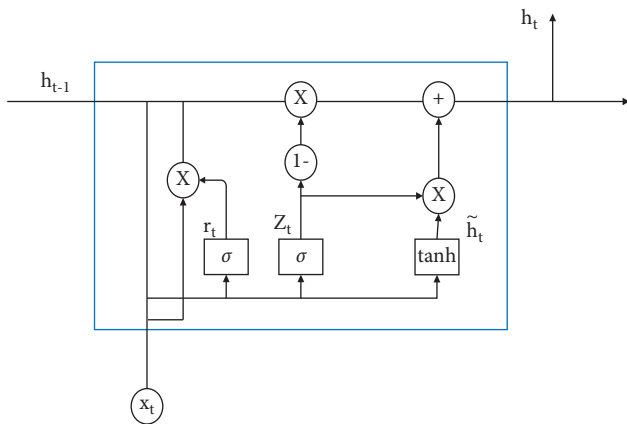


FIGURE 3: GRU network.

function,  $r_t$  represents the reset gate vector, and  $h_t$  represents the output vector.

**2.5.3. CNN Architecture.** CNNs were originally used to address image problems, but in natural language processing, the use of CNNs for text classification has achieved better results. In this paper, when the CNN is used for text classification, a six-layer convolution layer and a maximum pooling layer are used, as shown in Figure 5. The model adopts one-dimensional convolution. Without changing the width of the text sequence, the pooling layer uses the maximum pooling strategy to continuously extract important features from the text data through one-dimensional convolution and maximum pooling and then uses a pooling layer to combine the gathered text feature information and

input it into the fully connected layer. Finally, the category information of the classified text is output.

**2.5.4. Model Fusion.** Commonly used model fusion strategies include the averaging method, stacking method, and majority voting method. The main idea of the averaging method is to average the prediction results of each model and use the average value as the final prediction result. The stacking method uses model cross-validation, combining the features between models and training the newly combined features into a new model. Through this repeated feature stacking method, a strong classifier is finally obtained. The method of majority voting involves calculating the statistics of the classification results of each classifier. Among them, the classifier with the most votes divides the final data points in the corresponding category. This method is simple to calculate and easy to implement, and the classification effect is better than that of a single classification approach. The whole model fusion process is shown in Figure 6. The fusion steps of the entire model can be seen in Figure 6; first, the text data set is preprocessed, and the features are extracted and converted into a matrix  $(x_1, x_2, x_{m-1}, \text{ and } x_m)$ . Then, the text feature vector is input into the designed network architecture for parallel training, and each network architecture uses different optimization algorithms. A total of  $n$  models are calculated; then these  $n$  models are tested in parallel on the test sets to obtain the prediction results of each model; and finally, these  $k$  results are selected through the fusion strategy to select the final prediction result as the classification result of this mode.

Based on the above fusion strategy, the main idea is as follows: the total number of models trained in parallel is  $k$ ,

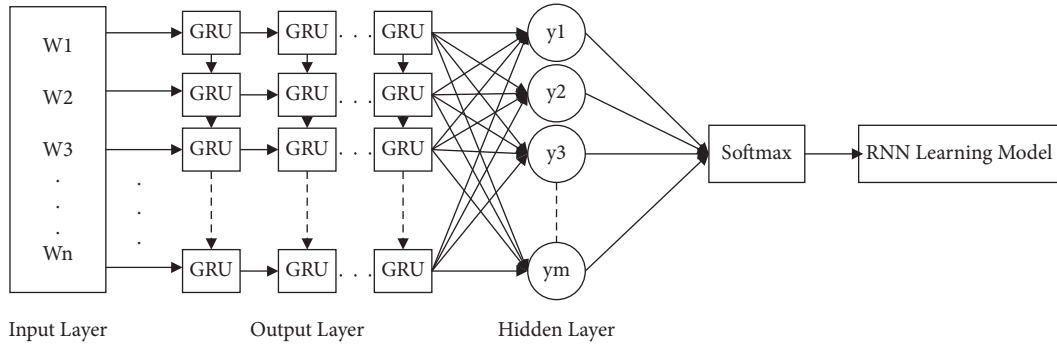


FIGURE 4: RNN architecture.

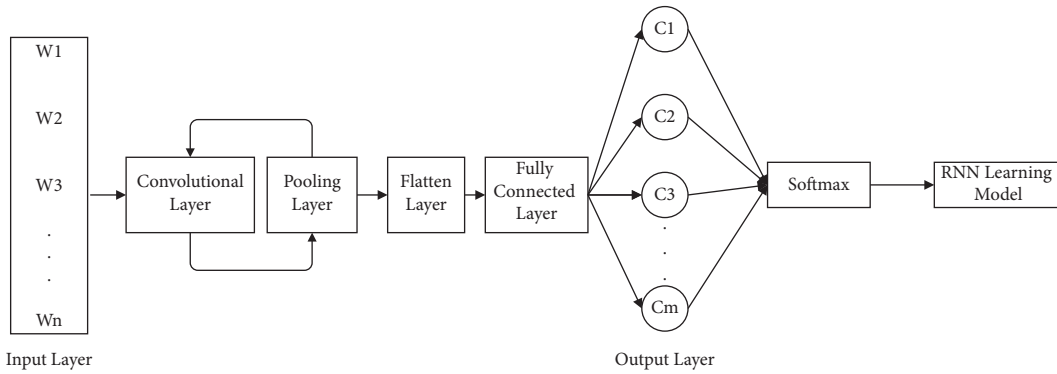


FIGURE 5: CNN architecture.

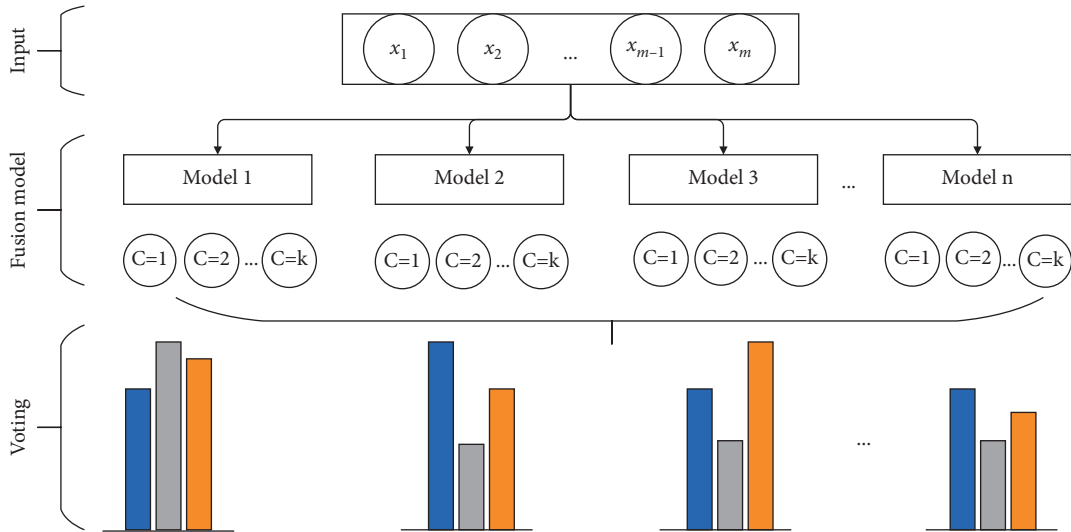


FIGURE 6: Model fusion structure.

the number of document categories is  $m$ , and the classification results of each model for text data  $i$  are counted. Among them, the text data  $i$  with the most votes is considered to belong to category  $c_{ij}$ . Finally, the accuracy rates belonging to category  $c_{ij}$  are summed and averaged as the final prediction result, and the calculations are shown in the following equations:

$$y_{ij} = [y_{i1}, y_{i2}, \dots, y_{ik}], \quad (23)$$

$$c_{ij} = \max[c_{i1}, c_{i2}, c_{i3}, \dots, c_{im}], \quad (24)$$

$$\hat{P}_{ij} = \frac{\sum_{n=1}^N \text{softmax}(y_{in})}{N}, \quad (25)$$

where  $y_{ij}$  represents the result of model  $j$  for the classification of text  $i$ ,  $k$  represents the number of training models,  $m$  represents the number of categories in the document,  $c_{im}$  represents the number of votes stating that text data  $i$  belongs to category  $m$ ,  $\hat{p}_{ij}$  represents the number of text data  $i$  belonging to the accuracy of category  $j$ , and  $N$  represents the number of votes.

### 3. Results and Discussion

**3.1. Datasets.** To verify that the fused model has generalizability, this paper uses three different data sets, as shown in Table 1: 20Newsgroups, Reuters, and Web of Science. The 20Newsgroups data set is mainly composed of 20 newsgroups for different topics. It contains 20,000 document data. The Reuters news data set contains 21,578 document data and a total of 90 categories. The Web of Science data set is a collection of academic article abstracts. The data sets include WOS5736 and WOS11967. This paper uses 80% of the documents as training data and 20% as test data. The relationships are shown in Table 1.

**3.2. Evaluation.** This paper uses accuracy, recall, and  $F_{1\text{-score}}$  to measure the classification performance of the model. The calculation equations are as follows:

$$\text{Accuracy} = \frac{(\text{TP} + \text{TN})}{(\text{TP} + \text{FP} + \text{FN} + \text{TN})}, \quad (26)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad (27)$$

$$F_{1\text{score}} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}}. \quad (28)$$

**3.3. Parameter Settings.** This paper uses the GloVe model to learn word vectors, as this model can realize global feature statistics in the corpus and local context feature vectors. At the same time, to prevent the overfitting of the training data, the dropout value is set to 0.5 in each network architecture. Each architecture model uses Adam, SGD, and RMSProp as three optimization algorithms and the ReLU activation function to improve the training speed of the model. The specific parameters are shown in Tables 2–4.

**3.4. Experimental Results.** To verify the classification performance of the proposed multimodal fusion approach and show that it is better than other single classification models, the experimental results obtained on three different public data sets are compared. The results are shown in Table 5.

Table 5 shows the classification results of the majority votes for nine models on the data sets. On the Reuters data set, the voting classification accuracy of the nine models reaches 89.23%, which is 0.66% higher than the model with the highest accuracy in the comparative experiments. The testing results of the fusion model are shown in Figure 7, and the comparative results of the four models are shown in

TABLE 1: Types of datasets.

Datasets	Training (%)	Testing (%)	Categories
20Newsgroups	80	20	20
Reuters	80	20	90
WOS5736	80	20	11
WOS11967	80	20	35

TABLE 2: DNN model parameters.

Parameter	Value
Word vector dimension	100
Learning rate	0.001
Hidden layers	6
Dropout	0.5
Activation function	ReLU
Hidden layer size	512
Batch size	64
Maximum number of words in text	500

TABLE 3: RNN model parameters.

Parameter	Value
Word vector dimension	100
Learning rate	0.001
Hidden layers	4
Dropout	0.5
Activation function	ReLU
Hidden layer size	64
Batch size	64
Maximum number of words in text	500

TABLE 4: CNN model parameters.

Parameter	Value
Word vector dimension	100
Learning rate	0.001
Convolutional layer	6
Flatten layer	1
Pooling layer	6
Fully connected layer	2
Activation function	ReLU
Dropout	0.5
Batch size	64
Maximum number of words in text	500

Figure 8. On the 20Newsgroups data set, when training the fusion model, 100 epochs are required for convergence, and the classification accuracy of the models reaches 88.87%, which is 5.13% higher than the model with the highest accuracy in the comparative experiments. The testing results of the fusion model and comparative models are shown in Figures 9 and 10, respectively. The classification accuracy rate that is achieved by the proposed models on the WOS5736 data set is 92.33%, which is 0.35% higher than the model with the highest accuracy in the comparative experiments. The testing results of the fusion model and comparative models are shown in Figures 11 and 12,

TABLE 5: Result comparison for text classification (unit %).

	Reuters			20Newsgroups			WOS5736			WOS11967		
	Acc	Recall	F1	Acc	Recall	F1	Acc	Recall	F1	Acc	Recall	F1
DNN	84.82	83.42	82.75	83.74	82.23	81.15	90.59	87.45	86.36	82.37	80.56	79.73
RNN	88.57	86.16	85.78	82.08	80.82	80.34	91.98	89.67	89.05	84.59	83.52	83.87
CNN	84.92	82.63	82.02	80.21	80.13	79.49	91.28	90.31	89.72	82.95	80.65	80.38
This paper	89.23	88.45	88.28	88.87	87.92	86.88	92.33	91.10	90.66	85.08	84.32	84.98

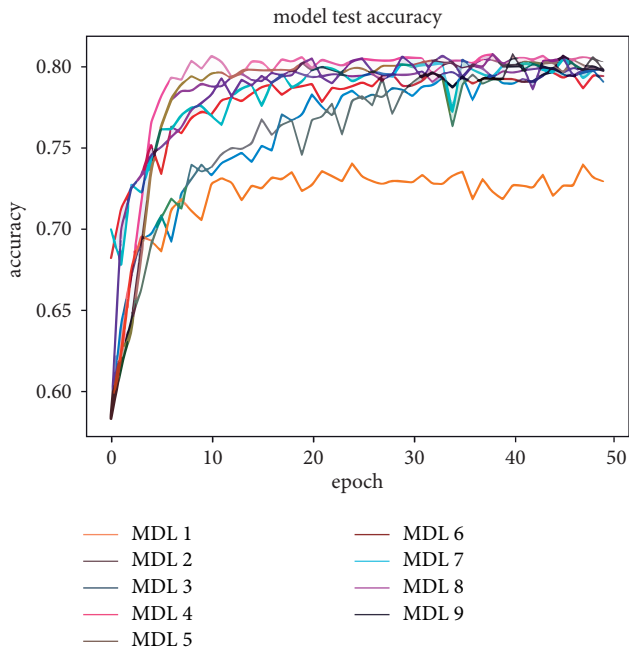


FIGURE 7: Reuters accuracy function for multimodel deep learning (MDL) trained by nine models in this paper.

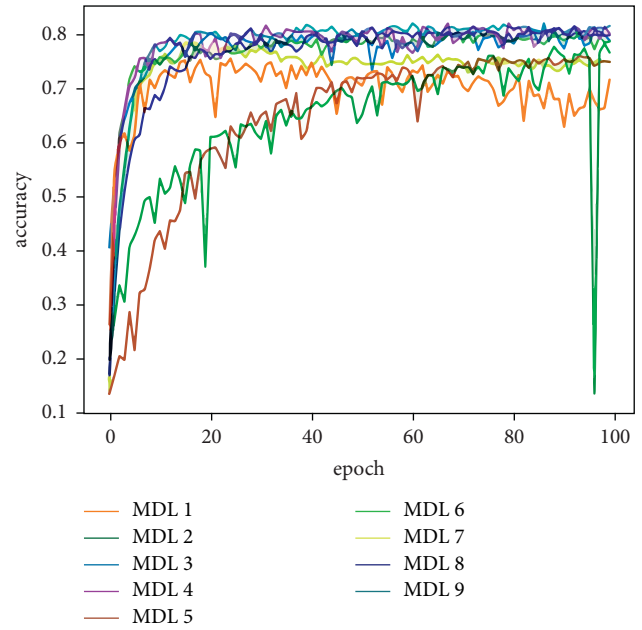


FIGURE 9: 20Newsgroups accuracy function for multimodel deep learning (MDL) trained by nine models in this paper.

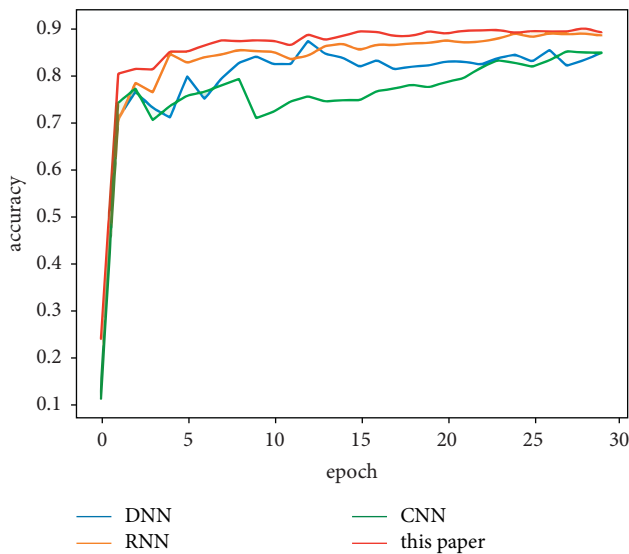


FIGURE 8: Reuters accuracy function for a comparative experiment in this paper.

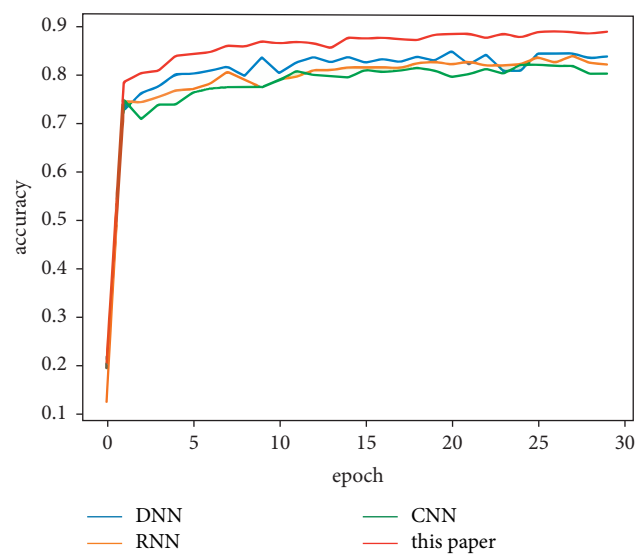


FIGURE 10: 20Newsgroups accuracy function for a comparative experiment in this paper.

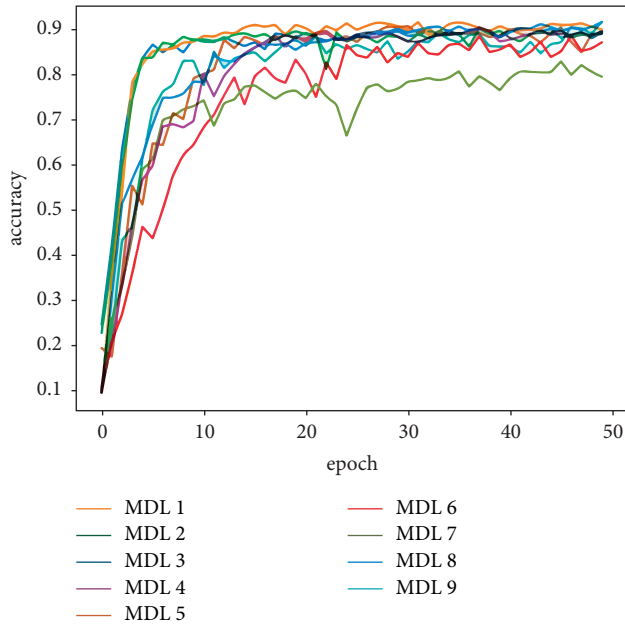


FIGURE 11: WOS5736 accuracy function for multimodel deep learning (MDL) trained by nine models in this paper.

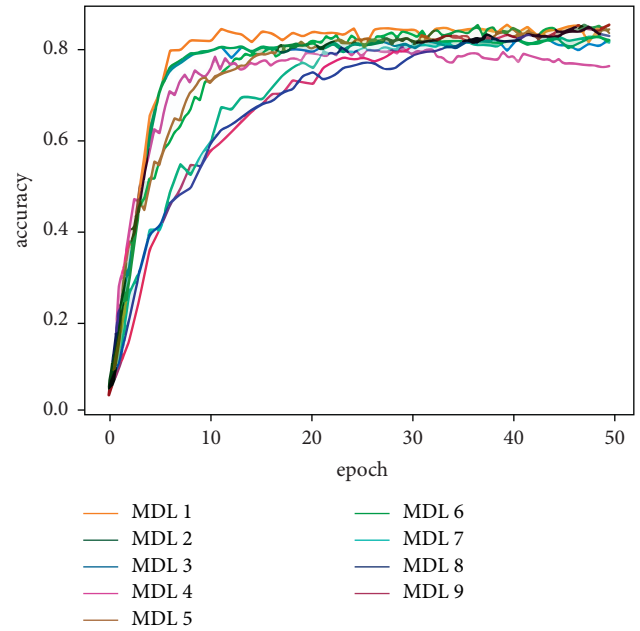


FIGURE 13: WOS11967 accuracy function for multimodel deep learning (MDL) trained by nine models in this paper.

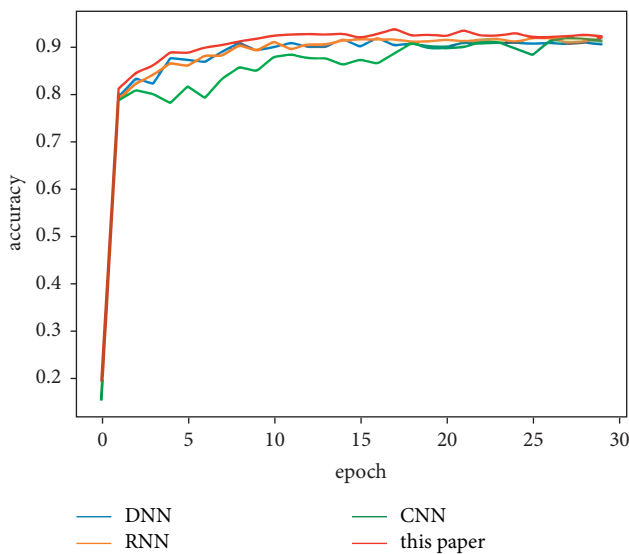


FIGURE 12: WOS5736 accuracy function for a comparative experiment in this paper.

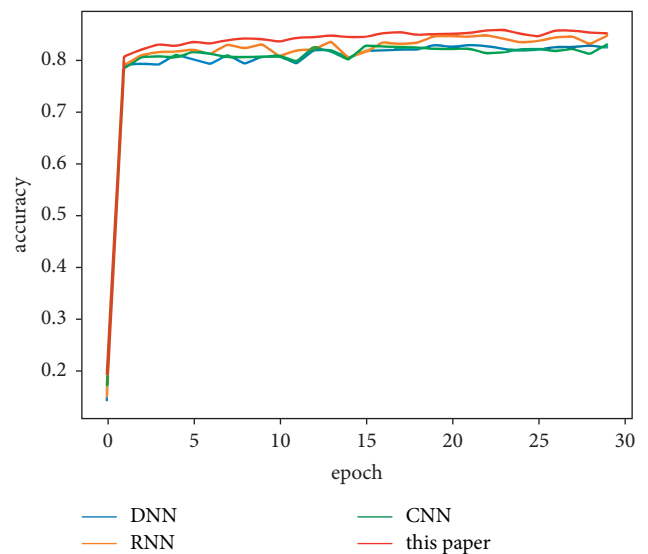


FIGURE 14: WOS11967 accuracy function for a comparative experiment in this paper.

respectively. The classification accuracy rate of the proposed models on the WOS11967 data set is 85.08%, which is 1.11% higher than the model with the highest accuracy in the comparative experiments. The testing results of the fusion model and comparative models are shown in Figures 13 and 14, respectively. Additionally, on four data sets, the recall and  $F1_{score}$  of the fusion model are higher than DNN, RNN, and CNN models. In summary, the model fusion method used in this paper has a better classification effect than a single classifier, and the fusion model has a more generalizable effect, which is specifically manifested as not only a better classification effect on a data set but also a good classification effect on a variety of data sets.

#### 4. Conclusion

A new classification method is proposed in this paper to solve the problems of data sets and the accuracy of a single model. The combination of parallel training of multiple deep learning architectures and integrated strategies is used to obtain the model. To verify the efficiency of the fusion model, the experimental evaluation of the fusion model on the Reuters, 20Newsgroups, WOS5736, and WOS11967 data sets shows that the accuracy, recall, and  $F1_{score}$  are higher using the fusion model compared with the DNN, RNN, and CNN models. The results show that the fusion model can

also improve text classification and an integration strategy can be used to provide flexibility for classification. This model also provides a new text classification method, which can be applied to a wide range of data sets. In future research, we will further explore the network structure of the fusion model and the influence of each network parameter on the classification results and analyze whether a shallow model can be used to improve the accuracy of the ultimate model while increasing training speed.

## Data Availability

Three types of data sets from different pages include the following. 20Newsgroups data set contains 20,000 documents with 20 categories. Reuters data set contains 21,578 documents with 90 categories. The Web of Science data set contains WOS11967 and WOS5736. WOS11967 contains 11,967 documents with 35 categories, which include 7 parents categories. WOS5736 contains 5,736 documents with 11 categories, which include 3 parents categories. These links are provided in these statements. All links are given below: (1) <https://archive.ics.uci.edu/ml/machine-learning-databases/20newsgroups-mld/>, (2) <http://kdd.ics.uci.edu/databases/reuters21578/reuters21578.html>, and (3) <https://data.mendeley.com/datasets/9rw3vkcfy4/2>.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This study was supported by the Natural Science Foundation of Shanxi Province, China (2021JQ-656).

## References

- [1] L. He, "Research and development of deep learning based text classification," *Computer Engineering*, vol. 47, no. 2, pp. 1–11, 2021, in Chinese.
- [2] J. Rocchio, "Relevance feedback in information retrieval," *The Smart retrieval system-experiments in automatic document processing*, Prentice Hall, Upper Saddle River, NJ, USA, 313–323, 1971.
- [3] B. J. Sowmya and K. G. Srinivasa, "Large scale multi-label text classification of a hierarchical dataset using rocchio algorithm," in *Proceedings of the 2016 International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS)*, pp. 291–296, Bengaluru, India, October 2016.
- [4] R. E. Schapire, "The strength of weak learnability," *Machine Learning*, vol. 5, no. 2, pp. 197–227, 1990.
- [5] L. Breiman, "Bagging predictors," *Machine Learning*, vol. 24, no. 2, pp. 123–140, 1996.
- [6] G. Ke, Z. Xu, J. Zhang, J. Bian, and T. Liu, "DeepGBM: a deep learning framework distilled by GBDT for online prediction tasks," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 384–394, Anchorage AK USA, August 2019.
- [7] S. Bloehdorn and A. Hotho, "Boosting for text classification with semantic features," in *Proceedings of the International Workshop on Knowledge Discovery on the Web*, pp. 149–166, Seattle, WA, USA, September 2017.
- [8] K. Kowsari, M. Heidarysafa, D. E. Brown, K. J. Meimandi, and L. E. Barnes, "Rmdl: random multimodel deep learning for classification," in *Proceedings of the 2nd International Conference on Information System and Data Mining*, pp. 19–28, Lakeland, FL, USA, April 2018.
- [9] S. Kaufmann, *CUBA: Artificial Conviviality and User-Behaviour Analysis in Web-Feeds*, University of Luxembourg, Luxembourg, Europe, 2010.
- [10] M. F. Porter, "An Algorithm for Suffix stripping," *Program*, vol. 14, 1980.
- [11] Y. Zhang, Y. Lan, Q. Fang, and X. J. Y. Xu, "Efficient reinforcement learning from demonstration via bayesian network-based knowledge extraction," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 7588221, 16 pages, 2021.
- [12] S. B. Kim, K. S. Kyoung-Soo Han, H. C. Hae-Chang Rim, and M. Sung Hyon, "Some effective techniques for naive Bayes text classification," *IEEE Transactions on Knowledge and Data Engineering*, vol. 18, no. 11, pp. 1457–1466, 2006.
- [13] H. Lodhi, C. Saunders, J. Shawe-Taylor, and N. Cristianini, "Text classification using string kernels," *Journal of Machine Learning Research*, vol. 2, pp. 419–444, 2002.
- [14] J. Chen, H. Jing, Y. Chang, and Q. Liu, "Gated recurrent unit based recurrent neural network for remaining useful life prediction of nonlinear deterioration process," *Reliability Engineering & System Safety*, vol. 185, pp. 372–382, 2019.
- [15] J. Schmidt-Hieber, "Nonparametric regression using deep neural networks with ReLU activation function," *Annals of Statistics*, vol. 48, no. 4, pp. 1875–1897, 2020.
- [16] L. C. Nguyen and H. Nguyen-Xuan, "Deep learning for computational structural optimization," *ISA Transactions*, vol. 103, pp. 177–191, 2020.
- [17] Y. HaCohen-Kerner, D. Miller, and Y. Yigal, "The influence of preprocessing on text classification using a bag-of-words representation," *PLoS One*, vol. 15, no. 5, Article ID e0232525, 2020.
- [18] D. Jatnika, M. A. Bijaksana, and A. A. Suryani, "Word2Vec model analysis for semantic similarities in English words," *Procedia Computer Science*, vol. 157, pp. 160–167, 2019.
- [19] F. Sakketou and N. Ampazis, "A constrained optimization algorithm for learning GloVe embeddings with semantic lexicons," *Knowledge-Based Systems*, vol. 195, Article ID 105628, 2020.
- [20] A. Jalilifard, V. F. Caridá, A. F. Mansano, R. S. Cristo, and F. P. C. da Fonseca, *Advances in Computing and Network Communications*, Springer, Singapore, pp. 327–337, 2021.
- [21] H. Wang, J. He, X. Zhang, and S. Liu, "A short text classification method based on N -gram and CNN," *Chinese Journal of Electronics*, vol. 29, no. 2, pp. 248–254, 2020.
- [22] A. Sherstinsky, "Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network," *Physica D: Nonlinear Phenomena*, vol. 404, Article ID 132306, 2020.



## Retraction

# Retracted: Tripartite Evolutionary Game Analysis of Financial Support for Science-Tech Enterprises' Innovation with Government Participation

### Computational Intelligence and Neuroscience

Received 1 August 2023; Accepted 1 August 2023; Published 2 August 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] W. Liu, Y. Liu, J. Gu, J. Zhao, and R. Zhang, "Tripartite Evolutionary Game Analysis of Financial Support for Science-Tech Enterprises' Innovation with Government Participation," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 5207003, 21 pages, 2022.

## Research Article

# Tripartite Evolutionary Game Analysis of Financial Support for Science-Tech Enterprises' Innovation with Government Participation

Wei Liu <sup>1</sup>, Yang Liu <sup>2</sup>, Jun Gu <sup>1</sup>, Jing Zhao <sup>1</sup>, and Rong Zhang <sup>1</sup>

<sup>1</sup>Business College of Beijing Union University, Beijing 100025, China

<sup>2</sup>Kunming University, Kunming, Yunnan 650214, China

Correspondence should be addressed to Rong Zhang; [rong.zhang@buu.edu.cn](mailto:rong.zhang@buu.edu.cn)

Received 4 February 2022; Accepted 31 March 2022; Published 30 April 2022

Academic Editor: Daqing Gong

Copyright © 2022 Wei Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this study, we use a tripartite evolutionary game model to analyze the financial support for science-tech enterprises' innovation with government participation. Based on the game model, the study uses replication dynamics to analyze the stability of evolutionary trajectory in the eight scenarios of enterprises, the government, and the financial institutes. Then, we discuss the effect of the reputation multiplier, the efficiency of firms' innovation output, and the positive socio-economic externalities multiplier by simulation as the important innovative points in this study. The study draws the following conclusions: (1) enterprises will choose high R&D intensity instead of low R&D intensity if the output of former exceeds that of the latter, (2) enterprises' R&D output is closely related to the strategy choices of the government and financial institutes, and (3) enterprises can attract government subsidy by boosting economic and social externalities. The government subsidies to enterprises with high R&D intensity will improve the innovation output by the innovation investment and reputation multiplier. So, the government can implement subsidy policy to boost high-intensive R&D activity. Financial institutes' strategy choice between equity investment and debt investment is influenced by investment yield difference and can influence enterprises' R&D intensity choices.

## 1. Introduction

Intensive innovation is an effective way for companies to improve their core competitiveness [1] (Frambach and Schillewaert) and is a core driver of high quality development. R&D intensity is the key to corporate innovation, and the strategic choice of R&D intensity by companies is mainly based on a cost-benefit analysis. The main source of funding for corporate innovation is internal and external financing, where external financing is constrained by the high risk of innovation projects, credit asymmetry, opportunism, and adverse selection [2, 3] (Brown et al.). Financial institutes disagree on firms' return on investment requirements [4] (Hall, 2002) and intensify financing constraints. Financial institutes' capital investment is considered beneficial to firms' R&D activities [5–7] (Fryges et al.), while higher financing costs arising from financing constraints reduce

firms' willingness to invest in R&D, thus reducing firms' willingness to innovate. Due to the positive externalities of technological innovation, governments are active promoters of firm innovation [8] (Yu et al.). A series of initiatives, including financial subsidies, can alleviate the financing constraints and stimulate firms to invest in innovation [9] (Zhou). On the one hand, government subsidies provide good investment signals for equity investment institutions [10–12] (Howell et al.), which is a disguised endorsement of firms' credit and R&D capabilities. On the contrary, government subsidies can reduce the cost of enterprise loan financing by the signal effect. Therefore, it is necessary for governments to understand how firms make choices about R&D intensity and how financial institutes make choices about loans or investment in R&D projects. Thus, government can influence the strategic choices of firms and financial institutes through subsidy policies.

Government subsidies provide the impetus for the development of enterprise innovation, enterprises are the subjects of innovation, and financial institutes provide financial services and financial support for enterprise innovation. Changes in the strategic choices of any of the participants in corporate innovation will lead to corresponding changes in stakeholder response strategies, both in terms of the continuous improvement of information structures and the adjustment of the rational level of multiple participants. There are a lot of research studies on enterprises' R&D strategy, financial institutes' investment strategy, and government subsidy policy. However, there is a lack of analysis that puts the three players in a comprehensive game framework. This study is based on this theoretical framework to introduce enterprises, financial institutes, and government into a comprehensive game in which enterprises choose high or low R&D intensity; financial institutes choose equity or debt investment, while the government chooses whether to provide subsidy. The government's subsidy policy, enterprises' R&D intensity choice, and financial institutions' investment structure influence each other. Therefore, the government can attract enterprises to choose high R&D intensity strategy and financial institutes to make equity investment by adjusting its subsidy policy.

The main innovations of this study are as follows:

- (1) Current research has focused on measuring risk-sharing ratios among financial institutes [13, 14] and lacks a holistic analysis of the behavior of governments, financial institutes and firms, and their interactions. However, examine the different strategic options and possible interactions between key players. Most previous research has only two players in the game framework without considering the function of government participating. The framework of this study is more complex with reality. The analysis focuses on the different strategic choices and possible interactions of firms, financial institutes, and governments in the process of obtaining external finance for firms with different R&D intensities. It draws some interesting conclusions.
- (2) Existing studies point out that government financial subsidies for business innovation vary across different types of enterprises [15]. The firms' R&D intensity has an impact on the effectiveness of government intervention. However, we find that the reputation multiplier of government subsidies also has the effect on firm innovation and the choice of investment approach by financial institutes. We also find that firms' innovation efficiency and strategy game affect the investment approach of financial institutes and the choice of government financial support, which in turn affects firms' financing costs.
- (3) Another possible contribution of this study is to explore how the efficiency of firms' innovation output and the positive socio-economic externalities generated by innovation impact on government and financial support decisions.

The remainder of this study is structured as follows. Section 2 briefly reviews the existing research on financial support for science-tech enterprises and evolutionary games and presents the literature gaps. Section 3 constructs a three-party evolutionary game model among SMEs, financial institutes, and government. Section 4 discusses the ESS of the tripartite game between the government, the financial institutes, and SMEs. Section 5 presents the numerical simulation result of the evolutionary game model, which discusses the effect of incentive-related parameters on the results. Section 6 provides discussion of evolutionary results and policy implications.

## 2. Literature Review

This study briefly reviews the literature on financial support for science-tech enterprises and evolutionary game models with the above objectives.

*2.1. Financial Support for Science-Tech Enterprises.* Through numerous studies, scholars have generally concluded that financing constraints have a negative impact on enterprises' R&D and innovation activities [16, 17] (Hottenrott et al.). Different scholars have conducted empirical studies using data from countries such as Germany, France, Spain, and Uruguay and have come to the consistent conclusion that financial support can be effective in stimulating enterprises' R&D investment [18–21] (Almus et al.). Some scholars studied the function of finance support to enterprise science-tech innovation through different channels. Abobal and Garda believed that technological innovation needs capital input such as bank loans [21]. Saint-Paul proposed financial development promotes technological progress by allowing economic entities to use technologies with greater risks but higher levels of productivity [22]. Luong et al. used firm-level data across 26 non-U.S. economies between 2000 and 2010 to show that foreign institutional ownership has a positive and causal effect on firm innovation [23]. Menezes and Pereira found financial product innovation speeds up technological innovation [15]. Kim et al. found that capital market financing could promote technological innovation of enterprises more than bank loans [24]. Research in China suggests that government and market forces need to work together to promote financial support for enterprises' innovation processes [25] (Gong et al.).

In the past few years, the effect of government subsidies on the external investors funding the enterprises' innovation has attracted significant scholarly attention. The government expects to stimulate R&D activities through direct subsidies. Sabrinaconduct that government subsidies have a positive effect on corporate patents [10]. Feldman & Kelley argue that companies that receive R&D subsidies are recognized by the government and send a signal that they are innovative and can attract more investment [26]. The idea is that companies that receive R&D grants are recognized by the government and send a signal that they are innovative, which can attract more investment [12, 27, 28] (Meuleman et al.). However,

the inevitable “rent-seeking behavior” can lead to resource misallocation due to government subsidies [29] (Lerner) and distortions in enterprises’ competitive behavior [30] (Guellec et al.) and distortions in the competitive behavior of enterprises. Chinese scholars have also demonstrated the positive effects of government subsidies on corporate R&D based on empirical evidence from China. For example, Liu et al. found that government subsidies stimulate enterprises’ R&D investment by using data on high-tech enterprises in Jiangsu [31].

**2.2. Evolutionary Game.** The central concepts of evolutionary game theory are evolutionary stable strategies and replicator dynamics, arguing that a group’s decisions can be achieved through dynamic behaviors such as imitation, communication, and learning between individuals, while having highly adaptive strategies. Strategies that are highly adaptive are more likely to be imitated by other participants; otherwise, these strategies are eliminated. The theory is an important mathematical tool because it is based on finite theory and finite information and is closer to reality.

In recent years, evolutionary games have rapidly developed into an active area of research in the socio-economic field and are an increasingly popular approach in the study of corporate innovation. Yang et al. developed an evolutionary game model between government, enterprises, universities, and research institutions to explore the mechanisms of intellectual property cooperation [32]. Han et al. used a game model to analyze the effect of cluster informal contracts on innovation cooperation among cluster enterprises and the impact of informal contracts on innovation cooperation among cluster enterprises [33]. Shen used evolutionary game theory to study firm decision-making behavior during open innovation [34]. The impact of heterogeneous structures on the evolution of innovation behavior has been discussed based on scale-free networks [35, 36] (Ma et al.; Su et al.). Some literature use the evolutionary game method to study the optimal government subsidy. For example, Meng and Zhao discussed the optimal subsidy that can promote the system to reach the ideal state by constructing the evolutionary game model between manufacturers and remanufacturers [37] (Meng and Zhao). Bi Peng and Chen used evolutionary game theory to study the cooperative innovation of key common technologies of the two equipment manufacturing enterprises, respectively, constructed an evolutionary game model with or without government macrocontrol, analyzed the key factors affecting the stability of system evolution, and pointed out the formulation of reasonable innovation income distribution scheme and effective regulation of enterprises by government departments [38]. Vaida believes that government policies and financial support are important factors affecting enterprise technological innovation [39]. Zhang constructs a tripartite evolutionary game model involving two manufacturing enterprises and the government in the monopolistic competitive market and analyzes the stability of each party’s strategy selection [40].

**2.3. Literature Gaps.** Most of the above studies have examined the impact of financing and government intervention on corporate technological innovation through an empirical approach, and these studies have clarified the impact of corporate financing on corporate technological innovation and the role of government in it. However, most of these studies have focused on the impact of government intervention or a single aspect of corporate financing constraints. The interaction between government, financial institutes, and enterprises is less often discussed. There is also little literature that explores in depth the different investment options of financial institutes. This study fills this gap by analyzing the interaction mechanism between government subsidies, financial institutes’ investment choices, and enterprises’ R&D investment intensity from a game perspective. Both equity financing and debt financing approaches are considered.

### 3. Research Methods

**3.1. Model Overview.** Science-tech enterprises, government, and financial institutes are very important stakeholders in innovation strategies. In this study, we will use the evolutionary game to study the influence of behavior interaction between the three game agents and analyze the evolutionary stability of the system under different circumstances and the important factors affecting its evolutionary stability.

The three-party game process is shown in Figure 1.

#### 3.2. Model Assumptions

**Assumption 1:** suppose the three game agents of science-tech enterprises, government, and financial institutes are limited rational agents with the goal of maximizing their own interests. They will constantly adjust and improve their strategic choices according to their own benefits in the game process.

**Assumption 2:** suppose the strategy of science-tech enterprises is “Innovation with High R&D Intensity” or “Innovation with Low R&D Intensity.” The government’s strategy is “Subsidy” or “no Subsidy,” and the strategy of financial institutes is “Equity investment” or “Debt investment.” The R&D expense of science-tech enterprises is all funding by the external financing from financial institutes and government.

**Assumption 3:** technological innovation can bring positive external economy and benefits to the government. The higher density of R & D investment in science-tech enterprises  $\alpha$ , the higher economic and social externality multiplier by innovation  $g$ , so we assume  $g_Q > g_L$ .

**Assumption 4:** reputation multiplier by government subsidies  $k(k > 1)$  enhance the output of enterprise innovation investment.

**Assumption 5:** government subsidies can share the risk of debt investment of financial institutes, so we assume  $r_S > r_N$ .

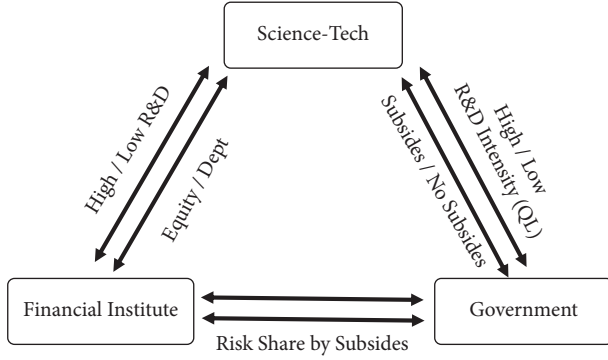


FIGURE 1: The three-party game process.

3.3. *Parameters' Setting.* The parameters' setting and explanations in the model are shown in Table 1.

3.4. *Model Construction.* Science-tech enterprise, government, and financial institute make strategic choices based on

$$\begin{aligned}
 E_Q &= yz(F_Q + S_Q)\alpha_Q\beta_Qk + (1-y)zF_Q\alpha_Q\beta_Q + y(1-z)[(F_Q + S_Q)\alpha_Q\beta_Qk - F_Q(r_S + 1)] \\
 &\quad + (1-y)(1-z)[F_Q\alpha_Q\beta_Q - F_Q(r_N + 1)], \\
 E_L &= yz(F_L + S_L)\alpha_L\beta_Lk + (1-y)zF_L\alpha_L\beta_L + y(1-z)[(F_L + S_L)\alpha_L\beta_Lk - F_L(r_S + 1)] \\
 &\quad + (1-y)(1-z)[F_L\alpha_L\beta_L - F_L(r_N + 1)], \\
 \bar{E}_I &= xE_Q + (1-x)E_L.
 \end{aligned} \tag{1}$$

4.1.2. *Government Expectation Function.* The expected return of government choosing "subsidy" is  $E_S$ , the expected

$$\begin{aligned}
 E_S &= x[(F_Q + S_Q)\alpha_Q\beta_Qkg_Q - S_Q - C_G] + (1-x)[(F_L + S_L)\alpha_L\beta_Lkg_L - S_L - C_G], \\
 E_N &= xF_Q\alpha_Q\beta_Qg_Q + (1-x)F_L\alpha_L\beta_Lg_L, \\
 \bar{E}_G &= yE_S + (1-y)E_N.
 \end{aligned} \tag{2}$$

4.1.3. *Financial Institute Expectation Function.* The expected return of financial institute choosing "equity investment" is

$$\begin{aligned}
 E_E &= xy(F_Q + S_Q)\alpha_Q\beta_Qk\delta + x(1-y)F_Q\alpha_Q\beta_Q\delta + (1-x)y(F_L + S_L)\alpha_L\beta_Lk\delta + (1-x)(1-y)F_L\alpha_L\beta_L\delta, \\
 E_D &= xyF_Qr_S + x(1-y)F_Qr_N + (1-x)yF_Lr_S + (1-x)(1-y)F_Lr_N, \\
 \bar{E}_F &= zE_E + (1-z)E_D.
 \end{aligned} \tag{3}$$

4.2. *Malthusian Replicated Dynamic Differential Equation Solution.* The replicated dynamic differential equation for the choice of an active strategy by science-tech enterprises,

self-willing. Assuming that the probability of science-tech enterprise choosing "innovation with high R & D intensity" strategy is  $x$ , the probability of choosing "innovation with low R & D intensity" strategy is  $1-x$ . The probability of the government choosing the "subsidy" strategy is  $y$ , and the probability of choosing the "no subsidy" strategy is  $1-y$ . If the probability of financial institutes choosing "equity investment" is  $z$ , the probability of choosing "debt investment" is  $1-z$ . Based on above, the income matrix of the three parties is shown in Table 2.

## 4. Results and Discussion

### 4.1. Construct Expectation Function

4.1.1. *Enterprise Expectation Function.* The expected return of enterprises' choosing strategy Q of "innovation with high R & D intensity" is  $E_Q$ , the expected return of enterprises choosing strategy L of "innovation with low R & D intensity" is  $E_L$ , and the average expected return  $\bar{E}_I$  will be

return of choosing "no subsidy" is  $E_N$ , and the average expected return  $\bar{E}_G$  will be

$E_E$ , the expected return of choosing "debt investment" is  $E_D$ , and the average expected return  $\bar{E}_F$  will be

government, and financial institutes can be expressed as follows:

TABLE 1: Parameter setting and explanations.

Parameters	Explanations
$F_Q$	Financial institutes' investments to enterprises with high R&D intensity
$F_L$	Financial institutes' investments to enterprises with low R&D intensity
$\alpha_Q$	R&D expense as % of financing for enterprises with high R&D intensity
$\alpha_L$	R&D expense as % of financing for enterprises with low R&D intensity
$\beta_Q$	Revenue per unit R&D expense for enterprises with high R&D intensity
$\beta_L$	Revenue per unit R&D expense for enterprises with low R&D intensity
$S_Q$	Government subsidies to enterprises with high R&D intensity
$S_L$	Government subsidies to enterprises with low R&D intensity
$k$	Reputation multiplier by government subsidies
$C_G$	Administrative costs of government subsidies
$g_Q$	Economic and social externality multiplier by innovation of enterprises with high R&D intensity
$g_L$	Economic and social externality multiplier by innovation of enterprises with low R&D intensity
$\delta$	Equity investment earnings as % of revenue
$r_S$	Debt investment yield with government subsidies
$r_N$	Debt investment yield without government subsidies
$x$	Probability of enterprises adopting high R&D intensity strategy
$y$	Probability of government adopting subsidy policy
$z$	Probability of equity investment from financial institutes

(1) Science-tech enterprises' replicated dynamic differential equation:

$$\begin{aligned}
F(x) &= \frac{dx}{dt} \\
&= x(E_Q - \bar{E}_I) \\
&= x(1-x)(E_Q - E_L) \\
&= x(1-x)[(1+yk-y)(\alpha_Q\beta_QF_Q - \alpha_L\beta_LF_L) + yk(\alpha_Q\beta_QS_Q - \alpha_L\beta_LS_L) - (1-z)(yr_S + r_N + 1 - yr_N)(F_Q - F_L)].
\end{aligned} \tag{4}$$

(2) Government's replicated dynamic differential equation:

$$\begin{aligned}
F(y) &= \frac{dy}{dt} \\
&= y(E_S - \bar{E}_G) \\
&= y(1-y)(E_S - E_N) \\
&= y(1-y)[x\alpha_Q\beta_Qg_Q(k-1)F_Q + x(\alpha_Q\beta_Qg_Qk-1)S_Q + (1-x)\alpha_L\beta_Lg_L(k-1)F_L + (1-z)(\alpha_L\beta_Lg_Lk-1)S_L - C_G].
\end{aligned} \tag{5}$$

(3) Financial institutes' replicated dynamic differential equation:

TABLE 2: Revenue matrix.

Strategy selection	Science-tech enterprises	Government	Equity investment ( $z$ )	Financial institute	Debt investment ( $1 - z$ )
Innovation with high/low R&D intensity	Q: high R&D intensity ( $x$ )	Subsidies ( $y$ )	$(F_Q + S_Q)\alpha_Q\beta_Q k(F_Q + S_Q)\alpha_Q\beta_Q k g_Q - S_Q - C_G(F_Q + S_Q)\alpha_Q\beta_Q k \delta$	$(F_Q + S_Q)\alpha_Q\beta_Q k - F_Q(\tau_s + 1)(F_Q + S_Q)\alpha_Q\beta_Q k g_Q - S_Q - C_G F_Q \tau_s$	
		No subsidies ( $1 - y$ )	$F_Q\alpha_Q\beta_Q F_Q\alpha_Q\beta_Q g_Q F_Q\alpha_Q\beta_Q \delta$	$F_Q\alpha_Q\beta_Q - F_Q(\tau_N + 1)F_Q\alpha_Q\beta_Q g_Q F_Q \tau_N$	
	L: low R&D intensity ( $1 - x$ )	Subsidies ( $y$ )	$(F_L + S_L)\alpha_L\beta_L k(F_L + S_L)\alpha_L\beta_L k g_L - S_L - C_G(F_L + S_L)\alpha_L\beta_L k \delta$	$(F_L + S_L)\alpha_L\beta_L k - F_L(\tau_s + 1)(F_L + S_L)\alpha_L\beta_L k g_L - S_L - C_G F_L \tau_s$	
		No subsidies ( $1 - y$ )	$F_L\alpha_L\beta_L F_L\alpha_L\beta_L g_L F_L\alpha_L\beta_L \delta$	$F_L\alpha_L\beta_L - F_L(\tau_N + 1)F_L\alpha_L\beta_L g_L F_L \tau_N$	

TABLE 3: Jacobian matrix eigenvalues.

Equalization points	Eigenvalues $\lambda_1$	Eigenvalues $\lambda_2$	Eigenvalues $\lambda_3$
$0(0, 0, 0)$	$\alpha_Q \beta_Q F_Q - \alpha_I \beta_I F_L - (r_N + 1)(F_Q - F_L)$	$\alpha_I \beta_I g_L (k - 1) F_L + (\alpha_I \beta_I g_I k - 1) S_L - C_G$	$(\alpha_I \beta_I \delta - r_N) F_L$
$A(0, 0, 1)$	$\alpha_Q \beta_Q F_Q - \alpha_I \beta_I F_L - \alpha_I \beta_I F_L$	$\alpha_I \beta_I g_L (k - 1) F_L + (\alpha_I \beta_I g_I k - 1) S_L - C_G$	$(r_N - \alpha_I \beta_I \delta) F_L$
$B(0, 1, 0)$	$k \alpha_Q \beta_Q (F_Q + S_Q) - k \alpha_I \beta_I (F_L + S_L) - (r_S + 1)(F_Q - F_L)$	$-\alpha_I \beta_I g_L (k - 1) F_L - (\alpha_I \beta_I g_I k - 1) S_L + C_G$	$k \alpha_I \beta_I \delta (F_L + S_L) - r_S F_L$
$C(0, 1, 1)$	$k \alpha_Q \beta_Q (F_Q + S_Q) - k \alpha_I \beta_I (F_L + S_L)$	$-\alpha_I \beta_I g_L (k - 1) F_L - (\alpha_I \beta_I g_I k - 1) S_L + C_G$	$-k \alpha_I \beta_I \delta (F_L + S_L) + r_S F_L$
$D(1, 0, 0)$	$\alpha_I \beta_I F_L - \alpha_Q \beta_Q F_Q + (r_N + 1)(F_Q - F_L)$	$\alpha_Q \beta_Q g_Q (k - 1) F_Q + (\alpha_Q \beta_Q g_Q k - 1) S_Q - C_G$	$(\alpha_Q \beta_Q \delta - r_N) F_Q$
$E(1, 0, 1)$	$\alpha_I \beta_I F_L - \alpha_Q \beta_Q F_Q$	$\alpha_Q \beta_Q g_Q (k - 1) F_Q + (\alpha_Q \beta_Q g_Q k - 1) S_Q - C_G$	$-(\alpha_Q \beta_Q \delta - r_N) F_Q$
$F(1, 1, 0)$	$k \alpha_I \beta_I (F_L + S_L) + (r_S + 1)(F_Q - F_L) - k \alpha_Q \beta_Q (F_Q + S_Q)$	$-\alpha_Q \beta_Q g_Q (k - 1) F_Q + (\alpha_Q \beta_Q g_Q k - 1) S_Q - C_G$	$k \alpha_Q \beta_Q \delta (F_Q + S_Q) - r_S F_Q$
$G(1, 1, 1)$	$k \alpha_I \beta_I (F_L + S_L) - k \alpha_Q \beta_Q (F_Q + S_Q)$	$-\alpha_Q \beta_Q g_Q (k - 1) F_Q + (\alpha_Q \beta_Q g_Q k - 1) S_Q - C_G$	$r_S F_Q - k \alpha_Q \beta_Q \delta (F_Q + S_Q)$



$$\begin{aligned}
F(z) &= \frac{dz}{dt} \\
&= z(E_E - \overline{E_F}) \\
&= z(1-z)(E_E - E_D) \\
&= z(1-z)\{x[(yk+1-y)\alpha_Q\beta_Q\delta - yr_S - (1-y)r_N]F_Q + xyk\alpha_Q\beta_Q\delta S_Q + (1-x)[(yk+1-y)\alpha_L\beta_L\delta \\
&\quad - yr_S - (1-y)r_N]F_L + (1-x)yk\alpha_L\beta_L\delta S_L\}.
\end{aligned} \tag{6}$$

Combining three equations, we will get the replication power system of science-tech enterprise, government, and financial institutes as below:

$$\begin{cases}
F(x) = x(1-x)[(1+yk-y)(\alpha_Q\beta_Q F_Q - \alpha_L\beta_L S_L) - (1-z)(yr_S + r_N + 1 - yr_N)(F_Q - F_L)], \\
F(y) = y(1-y)[x\alpha_Q\beta_Q g_Q(k-1)F_Q + x(\alpha_Q\beta_Q g_Q k - 1)S_Q + (1-x)\alpha_L\beta_L g_L(k-1)F_L + (1-z)(\alpha_L\beta_L g_L k - 1)S_L - C_G], \\
F(z) = z(1-z)\{x[(yk+1-y)\alpha_Q\beta_Q\delta - yr_S - (1-y)r_N]F_Q + xyk\alpha_Q\beta_Q\delta S_Q + (1-x)[(yk+1-y)\alpha_L\beta_L\delta - yr_S - (1-y)r_N] \\
F_L + (1-x)yk\alpha_L\beta_L\delta S_L\}.
\end{cases} \tag{7}$$

### 4.3. Strategy Portfolio Stability Analysis

4.3.1. *Equilibrium of Dynamic System.* According to Takalo and Tanayama [11], we can judge the strategy portfolio stability by analyzing the Jacobian matrix; then, we conduct the replicated dynamic system Jacobian matrix:

$$J = \begin{pmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{pmatrix}. \tag{8}$$

Among them,

$$P_{11} = \frac{\partial F(x)}{\partial x}$$

$$= (1 - 2x)[(1 + yk - y)(\alpha_Q\beta_Q F_Q - \alpha_L\beta_L F_L) + yk(\alpha_Q\beta_Q S_Q - \alpha_L\beta_L S_L) - (1 - z)(yr_S + r_N + 1 - yr_N)(F_Q - F_L)],$$

$$P_{12} = \frac{\partial F(x)}{\partial y}$$

$$= x(1 - x)[(k - 1)(\alpha_Q\beta_Q F_Q - \alpha_L\beta_L F_L) + k(\alpha_Q\beta_Q S_Q - \alpha_L\beta_L S_L) - (1 - z)(r_S - r_N)(F_Q - F_L)],$$

$$P_{13} = \frac{\partial F(x)}{\partial z}$$

$$= x(1 - x)\left(yr_S + \text{tarts} - \text{with}(., 'hskip')\right)''hskip''\text{substring} - \text{after}(\text{preceding} - \text{sibling} :: \text{comment}()[\text{starts} - \text{with}(., 'hskip')], 'hskip')$$

(9)

In the dynamic system composed of three game agents, assuming  $F(x) = 0$ ,  $F(y) = 0$ , and  $F(z) = 0$ , we obtain eight Nash equilibrium points of the system: O(0, 0, 0), A(0, 0, 1), B(0, 1, 0), C(0, 1, 1), D(1, 0, 0), E(1, 0, 1), F(1, 1, 0), and G(1, 1, 1).

According to Lyapunov's first law, if the eigenvalues of the corresponding matrix are all negative, the equilibrium point is the system's ESS. The stability of each point is shown in Table 3.

**4.3.2. Equilibrium Stability Analysis.** The calculations reveal that this study should analyze the game's stabilization strategy in 8 scenarios.

Scenario 1: if the following conditions are satisfied, the equilibrium is G(1, 1, 1):

$$\begin{cases} \alpha_L\beta_L(F_L + S_L) < \alpha_Q\beta_Q(F_Q + S_Q), \\ C_G < \alpha_Q\beta_Q g_Q(k - 1)F_Q + (\alpha_Q\beta_Q g_Q k - 1)S_Q, \\ r_S F_Q < k\alpha_Q\beta_Q \delta(S_Q + F_Q). \end{cases} \quad (10)$$

When  $\alpha_L\beta_L(F_L + S_L) < \alpha_Q\beta_Q(F_Q + S_Q)$ ,  $C_G < \alpha_Q\beta_Q g_Q(k - 1)F_Q + (\alpha_Q\beta_Q g_Q k - 1)S_Q$ ,  $r_S F_Q < k\alpha_Q\beta_Q \delta(S_Q + F_Q)$ , the eigenvalues of the Jacobi matrix corresponding to the evolutionary stability point G(1, 1, 1) are negative, indicating that G(1, 1, 1) is ESS. This indicates that science-tech enterprises choose high R&D intensity when the benefits from innovation investment of high R&D intensity firms exceeds that of low R&D intensity firms. When the socio-economic benefits increased by government subsidies outweigh the administrative costs, the government chooses to implement subsidies for high R&D-intensive enterprises. Therefore, financial institutes choose to make equity investments when the return on equity investment in high R&D intensity enterprises that receive government subsidies is greater than the return on debt investment.

Scenario 2: if the following conditions are satisfied, the equilibrium is F(1, 1, 0):

$$\begin{cases} k\alpha_L\beta_L(F_L + S_L) + (r_S + 1)(F_Q - F_L) < k\alpha_Q\beta_Q(F_Q + S_Q), \\ C_G < \alpha_Q\beta_Q g_Q(k - 1)F_Q + (\alpha_Q\beta_Q g_Q k - 1)S_Q, \\ k\alpha_Q\beta_Q \delta(S_Q + F_Q) < r_S F_Q. \end{cases} \quad (11)$$

When  $k\alpha_L\beta_L(F_L + S_L) + (r_S + 1)(F_Q - F_L) < k\alpha_Q\beta_Q(F_Q + S_Q)$ ,  $C_G < \alpha_Q\beta_Q g_Q(k - 1)F_Q + (\alpha_Q\beta_Q g_Q k - 1)S_Q$ , and  $k\alpha_Q\beta_Q \delta(S_Q + F_Q) < r_S F_Q$ , the eigenvalues of the Jacobi matrix corresponding to the evolutionary stability point F(1, 1, 0) are negative, indicating that F(1, 1, 0) is ESS. This indicates that enterprises choose a high R&D intensity when the total benefits termed by the high R&D intensity technology are higher than those received by the low R&D intensity firm. The total benefits include the subsidies received from the government and the resulting reputational value added (this reputational value added mainly comes in the form of obtaining lower debt financing costs). The government will choose to subsidize high R&D intensity enterprises when the socio-economic benefits of government subsidies outweigh the administrative costs. Financial institutes, on the contrary, have a greater return on debt investment than on equity investment and therefore choose debt investment.

Scenario 3: if the following conditions are satisfied, the equilibrium is E(1, 0, 1):

$$\begin{cases} \alpha_L\beta_L F_L < \alpha_Q\beta_Q F_Q, \\ \alpha_Q\beta_Q g_Q(k - 1)F_Q + (\alpha_Q\beta_Q g_Q k - 1)S_Q < C_G, \\ r_N < \alpha_Q\beta_Q \delta. \end{cases} \quad (12)$$

When  $\alpha_L\beta_L F_L < \alpha_Q\beta_Q F_Q$ ,  $\alpha_Q\beta_Q g_Q(k - 1)F_Q + (\alpha_Q\beta_Q g_Q k - 1)S_Q < C_G$ , and  $r_N < \alpha_Q\beta_Q \delta$ , the eigenvalues of

the Jacobi matrix corresponding to the evolutionary stability point  $E(1, 0, 1)$  are negative, indicating that  $E(1, 0, 1)$  is ESS. This indicates that enterprises that receive equity investment from financial institutes for high R&D intensity will choose high R&D intensity when the return is greater than the return for low R&D intensity. The government will not subsidize when the economic and social benefits gained from government subsidies are not sufficient to offset the cost of administration. Financial institutes choose to make equity investments in high R&D intensity enterprises when the rate of return on equity investments in science-tech enterprises is greater than that of return on debt investments.

Scenario 4: if satisfy the following conditions, the equilibrium is D  $(1, 0, 0)$ :

$$\begin{cases} \alpha_L \beta_L F_L + (r_N + 1)(F_Q - F_L) < \alpha_Q \beta_Q F_Q, \\ \alpha_Q \beta_Q g_Q (k-1)F_Q + (\alpha_Q \beta_Q g_Q k - 1)S_Q < C_G, \\ \alpha_Q \beta_Q \delta < r_N. \end{cases} \quad (13)$$

When  $\alpha_L \beta_L F_L + (r_N + 1)(F_Q - F_L) < \alpha_Q \beta_Q F_Q$ ,  $\alpha_Q \beta_Q g_Q (k-1)F_Q + (\alpha_Q \beta_Q g_Q k - 1)S_Q < C_G$ , and  $\alpha_Q \beta_Q \delta < r_N$ , the eigenvalues of the Jacobi matrix corresponding to the evolutionary stability point D  $(1, 0, 0)$  are negative, indicating that D  $(1, 0, 0)$  is ESS. This indicates that enterprises that receive loans from financial institutes will choose high R&D intensity when the return of innovation investment for high R&D intensity is greater than the return for low R&D intensity. The government subsidizes when the economic and social benefits gained from government subsidies are not sufficient to offset the cost of administration. Financial institutes choose to make debt investments in high R&D intensity enterprises when the return on equity investment in science-tech enterprises is less than the return on debt investment.

Scenario 5: if the following conditions are satisfied, the equilibrium is C  $(0, 1, 1)$ :

$$\begin{cases} \alpha_Q \beta_Q (F_Q + S_Q) < \alpha_L \beta_L (F_L + S_L), \\ C_G < \alpha_L \beta_L g_L (k-1)F_L + (\alpha_L \beta_L g_L k - 1)S_L, \\ r_S F_L < k \alpha_L \beta_L \delta (S_L + F_L). \end{cases} \quad (14)$$

When  $\alpha_Q \beta_Q (F_Q + S_Q) < \alpha_L \beta_L (F_L + S_L)$ ,  $C_G < \alpha_L \beta_L g_L (k-1)F_L + (\alpha_L \beta_L g_L k - 1)S_L$ , and  $r_S F_L < k \alpha_L \beta_L \delta (S_L + F_L)$ , the eigenvalues of the Jacobi matrix corresponding to the evolutionary stability point C  $(0, 1, 1)$  are negative, indicating that C  $(0, 1, 1)$  is ESS. This indicates that the enterprises choose to make low R&D intensity, when the benefits of high R&D intensity are less than the benefits of low R&D intensity. The government chooses to subsidize when the economic and social benefits gained from government subsidies outweigh the administrative costs. Financial institutes choose to make equity investments in low R&D

intensity enterprises when the return on equity investment in low R&D intensity science-tech enterprises is greater than the return on debt investment. Science-tech enterprises receive government subsidies and equity financing to invest in R&D.

Scenario 6: if the following conditions are satisfied, the equilibrium is B  $(0, 1, 0)$ :

$$\begin{cases} k \alpha_Q \beta_Q (F_Q + S_Q) - (r_S + 1)(F_Q - F_L) < k \alpha_L \beta_L (F_L + S_L), \\ C_G < \alpha_L \beta_L g_L (k-1)F_L + (\alpha_L \beta_L g_L k - 1)S_L, \\ k \alpha_L \beta_L \delta (S_L + F_L) < r_S F_L. \end{cases} \quad (15)$$

When  $k \alpha_Q \beta_Q (F_Q + S_Q) - (r_S + 1)(F_Q - F_L) < k \alpha_L \beta_L (F_L + S_L)$ ,  $C_G < \alpha_L \beta_L g_L (k-1)F_L + (\alpha_L \beta_L g_L k - 1)S_L$ , and  $k \alpha_L \beta_L \delta (S_L + F_L) < r_S F_L$ , the eigenvalues of the Jacobi matrix corresponding to the evolutionary stability point B  $(0, 1, 0)$  are negative, indicating that B  $(0, 1, 0)$  is ESS. This indicates that the enterprises choose low R&D intensity, when the benefits of high R&D intensity are less than the benefits obtained from low R&D intensity. The government chooses to subsidize when the economic and social benefits gained from government subsidies outweigh the administrative costs. Financial institutes choose to make debt investments in low R&D intensity enterprises when the return on equity investment is less than the return on debt investment. Science-tech enterprises receive government subsidies and debt financing for R&D investment.

Scenario 7: if the following conditions are satisfied, the equilibrium is A  $(0, 0, 1)$ :

$$\begin{cases} \alpha_Q \beta_Q F_Q < \alpha_L \beta_L F_L, \\ \alpha_L \beta_L g_L (k-1)F_L + (\alpha_L \beta_L g_L k - 1)S_L < C_G, \\ r_N < \alpha_L \beta_L \delta. \end{cases} \quad (16)$$

When  $\alpha_Q \beta_Q F_Q < \alpha_L \beta_L F_L$ ,  $\alpha_L \beta_L g_L (k-1)F_L + (\alpha_L \beta_L g_L k - 1)S_L < C_G$ , and  $r_N < \alpha_L \beta_L \delta$ , the eigenvalues of the Jacobi matrix corresponding to the evolutionary stability point A  $(0, 0, 1)$  are negative, indicating that A  $(0, 0, 1)$  is ESS. This indicates that enterprises choose low R&D intensity when the return of innovation investment is less. The government will not subsidize when the economic and social benefits gained are not sufficient to offset the cost of administration. Financial institutes choose to make equity investments in low R&D intensity enterprises when the return is greater than that on debt investment.

Scenario 8: if the following conditions are satisfied, the equilibrium is O  $(0, 0, 0)$ :

$$\begin{cases} \alpha_Q \beta_Q F_Q - (r_N + 1)(F_Q - F_L) < \alpha_L \beta_L F_L, \\ \alpha_L \beta_L g_L (k-1)F_L + (\alpha_L \beta_L g_L k - 1)S_L < C_G, \\ \alpha_L \beta_L \delta < r_N. \end{cases} \quad (17)$$

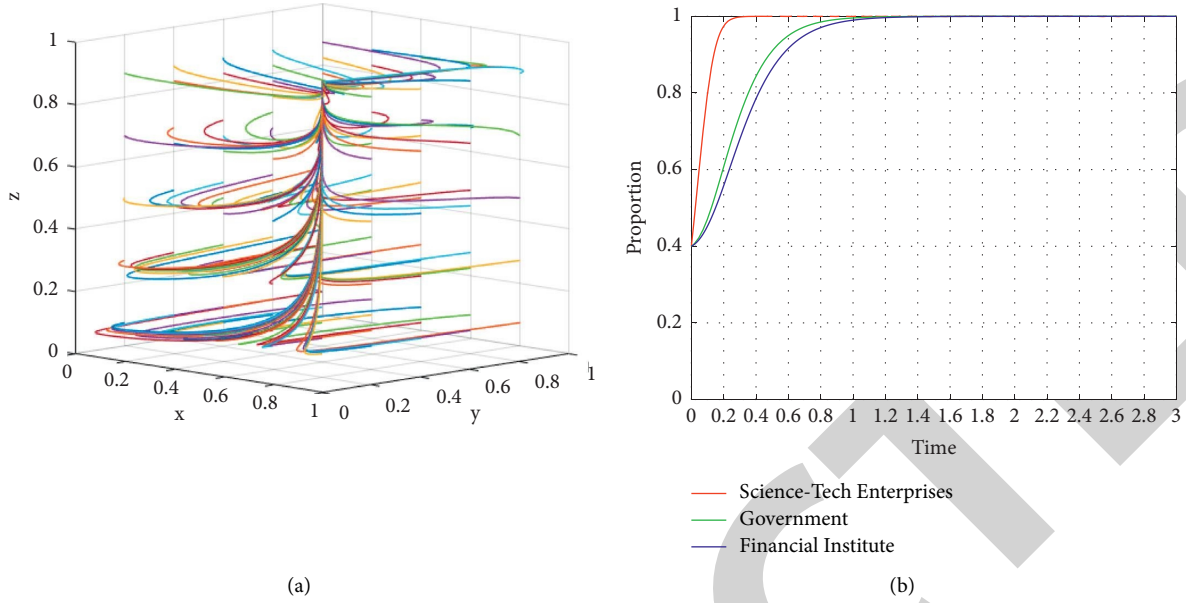


FIGURE 2: (a) Three-dimensional situation of evolutionary trajectory of  $G(1, 1, 1)$ . (b) Two-dimensional situation of evolutionary trajectory of  $G(1, 1, 1)$ .

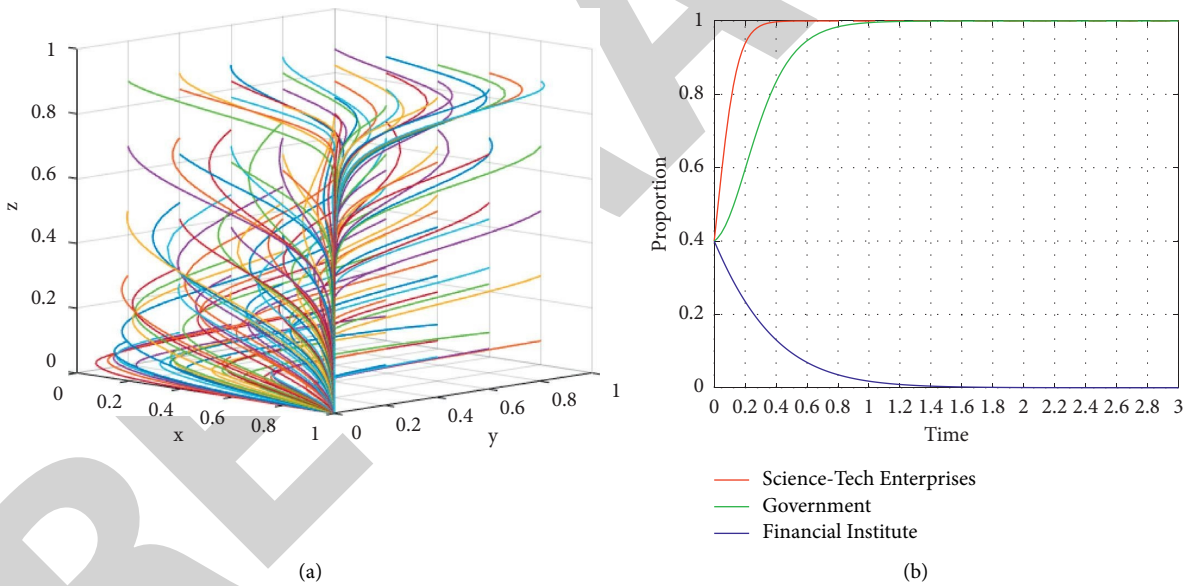


FIGURE 3: (a) Three-dimensional situation of evolutionary trajectory of  $F(1, 1, 0)$ . (b) Two-dimensional situation of evolutionary trajectory of  $F(1, 1, 0)$ .

When  $\alpha_Q \beta_Q F_Q - (r_N + 1)(F_Q - F_L) < \alpha_L \beta_L F_L$ ,  $\alpha_L \beta_L g_L (k - 1)F_L + (\alpha_L \beta_L g_L k - 1)S_L < C_G$ , and  $\alpha_L \beta_L \delta < r_N$ , the eigenvalues of the Jacobi matrix corresponding to the evolutionary stability point  $O(0, 0, 0)$  are negative, indicating that  $O(0, 0, 0)$  is ESS. This indicates that enterprises will choose low R&D intensity, when the profit of innovation investment is less. The government will not subsidize when the economic and social benefits gained from government subsidies are not sufficient to offset the cost of administration. Financial institutes choose debt investments in low

R&D intensity enterprises when the return on equity investments is less than the return on debt investments.

Based on the above evolutionary game model, we analyze the eight scenarios of different strategy portfolios chosen by enterprises, government, and financial institutes. For the three players, one's choice is the condition which can influence other two's choices. Enterprises choose high R&D intensity instead of low R&D intensity if the output of former exceeds that of the latter. Enterprises' R&D output is

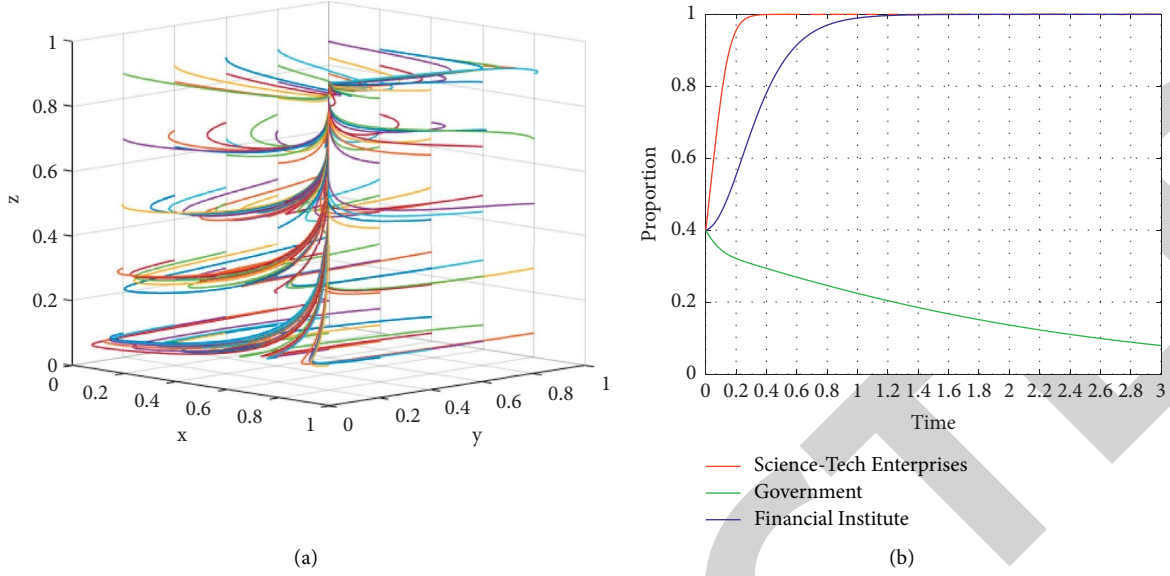


FIGURE 4: (a) Three-dimensional situation of evolutionary trajectory of E (1, 0, 1). (b) Two-dimensional situation of evolutionary trajectory of E (1, 0, 1).

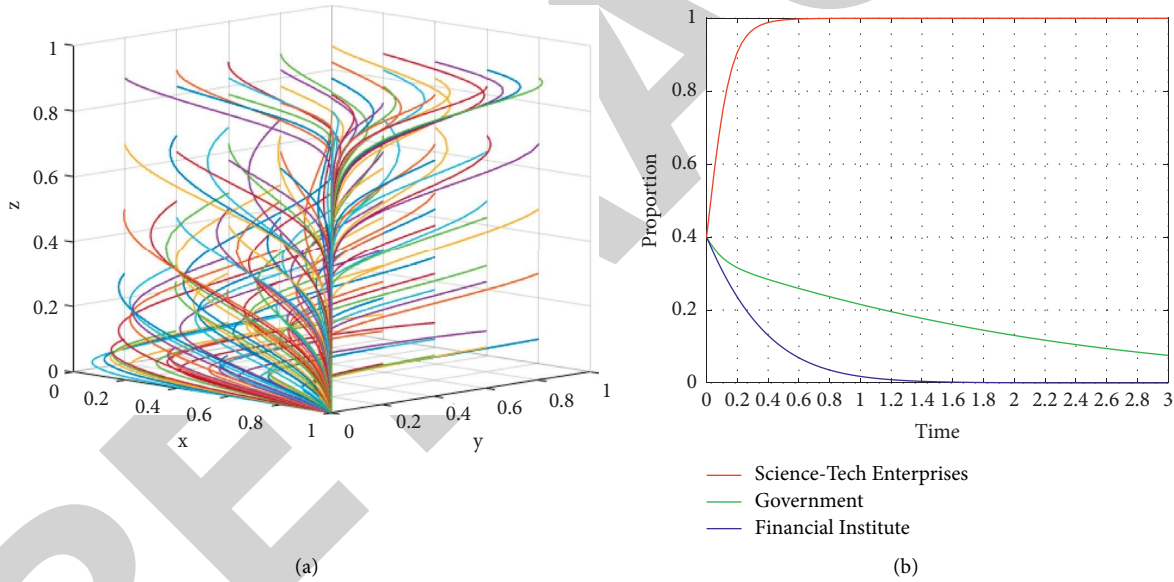


FIGURE 5: (a) Three-dimensional situation of evolutionary trajectory of D (1, 0, 0). (b) Two-dimensional situation of evolutionary trajectory of D (1, 0, 0).

closely related to the strategy choices of the government and financial institutes. The government chooses provide subsidy if its benefits exceed its costs. As the government pays attention to broad economic and social targets, the benefits include not only the R&D output itself but also economic and social externalities and the regulation effect to enterprises. The cost includes subsidies as well as administrative costs. Financial institutes make strategy choices between equity investment and debt investment based on their investment yields. The debt investment yield is different with government subsidies from that without government subsidies.

## 5. Numerical Simulation

**5.1. The Evolutionary Trajectory of ESS.** For analyzing the dynamic evolution process, the strategy evolution process of the tripartite game in different scenarios can be simulated by changing the parameter settings. In this section, different stability conditions are brought into MATLAB R2022a to simulate the evolutionary trajectory.

- (1) Assume  $F_Q = 100$ ,  $F_L = 80$ ,  $\alpha_Q = 0.3$ ,  $\alpha_L = 0.2$ ,  $\beta_Q = 1.5$ ,  $\beta_L = 1.2$ ,  $S_Q = 10$ ,  $S_L = 8$ ,  $k = 1.2$ ,  $C_G = 1$ ,  $g_Q = 1.2$ ,  $g_L = 1.1$ ,  $\delta = 0.3$ ,  $r_S = 0.05$ , and  $r_N = 0.08$ . The evolutionary trajectory of G (1, 1, 1) is featured in



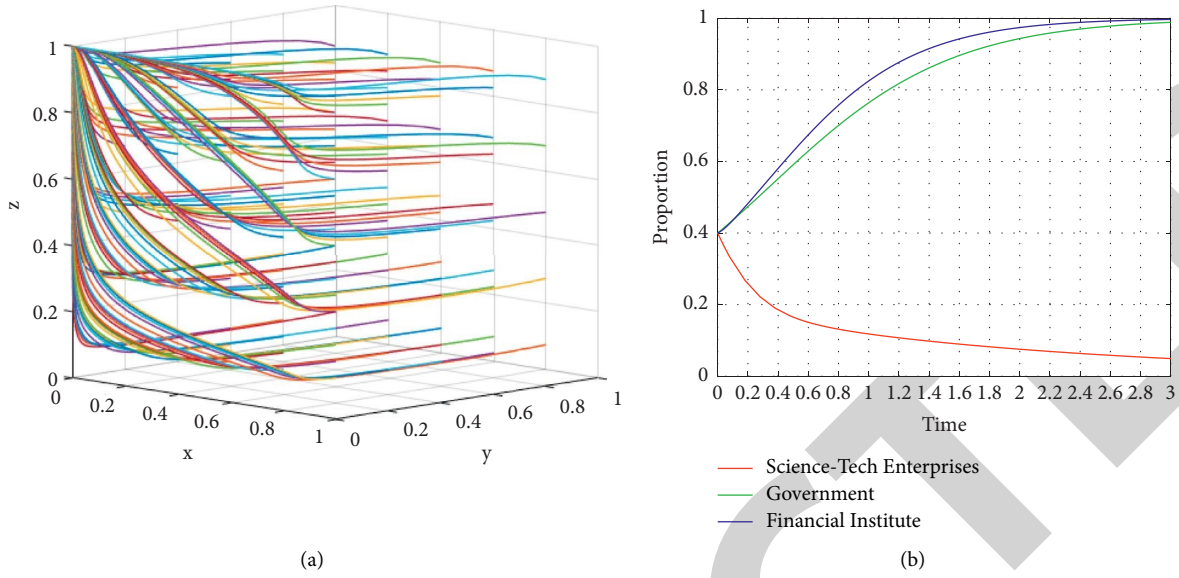


FIGURE 6: (a) Three-dimensional situation of evolutionary trajectory of  $C(0, 1, 1)$ . (b) Two-dimensional situation of evolutionary trajectory of  $C(0, 1, 1)$ .

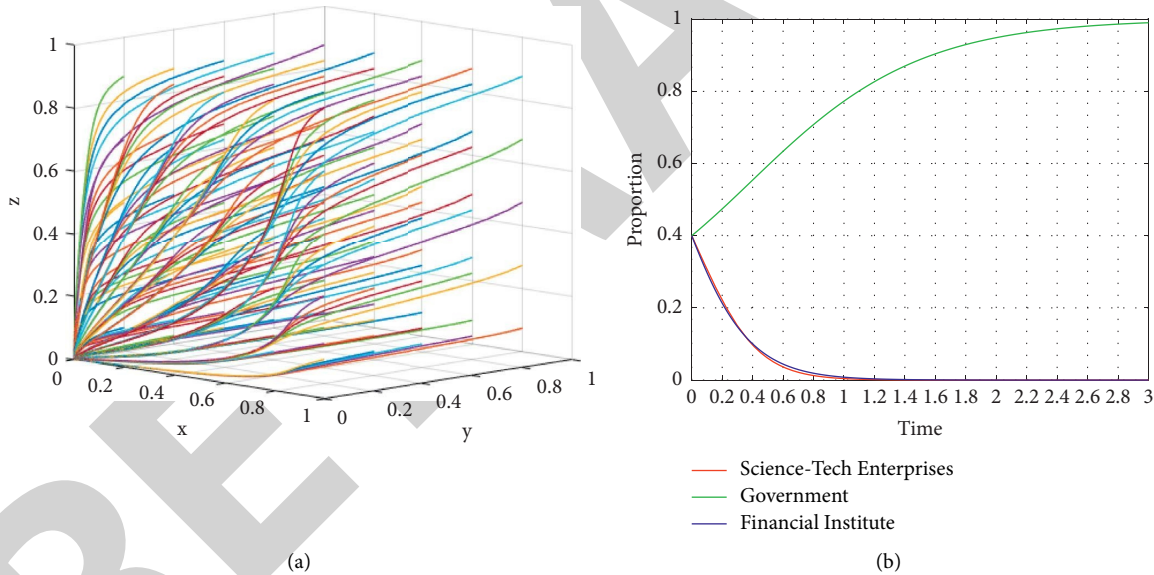


FIGURE 7: (a) Three-dimensional situation of evolutionary trajectory of  $B(0, 1, 0)$ . (b) Two-dimensional situation of evolutionary trajectory of  $B(0, 1, 0)$ .

Figure 2(a). When the initial probabilities of all three parties are 0.4, the evolutionary trajectory is displayed in Figure 2(b).

- (2) Assume  $F_Q = 100$ ,  $F_L = 80$ ,  $\alpha_Q = 0.3$ ,  $\alpha_L = 0.2$ ,  $\beta_Q = 1.5$ ,  $\beta_L = 1.2$ ,  $S_Q = 10$ ,  $S_L = 8$ ,  $k = 1.2$ ,  $C_G = 1$ ,  $g_Q = 1.2$ ,  $g_L = 1.1$ ,  $\delta = 0.1$ ,  $r_S = 0.06$ , and  $r_N = 0.08$ . The evolutionary trajectory of  $F(1, 1, 0)$  is featured in Figure 3(a). When the initial probabilities of all three parties are 0.4, the evolutionary trajectory is displayed in Figure 3(b).

- (3) Assume  $F_Q = 100$ ,  $F_L = 80$ ,  $\alpha_Q = 0.3$ ,  $\alpha_L = 0.2$ ,  $\beta_Q = 1.5$ ,  $\beta_L = 1.2$ ,  $S_Q = 10$ ,  $S_L = 8$ ,  $k = 1.1$ ,  $C_G = 1$ ,  $g_Q = 1.1$ ,  $g_L = 1.05$ ,  $\delta = 0.3$ ,  $r_S = 0.06$ , and  $r_N = 0.08$ . The evolutionary trajectory of  $E(1, 0, 1)$  is featured in Figure 4(a). When the initial probabilities of all three parties are 0.4, the evolutionary trajectory is displayed in Figure 4(b).

- (4) Assume  $F_Q = 100$ ,  $F_L = 80$ ,  $\alpha_Q = 0.3$ ,  $\alpha_L = 0.2$ ,  $\beta_Q = 1.5$ ,  $\beta_L = 1.2$ ,  $S_Q = 10$ ,  $S_L = 8$ ,  $k = 1.1$ ,  $C_G = 1$ ,  $g_Q = 1.1$ ,  $g_L = 1.05$ ,  $\delta = 0.1$ ,  $r_S = 0.06$ , and  $r_N = 0.08$ . The evolutionary trajectory of  $D(1, 0, 0)$  is featured

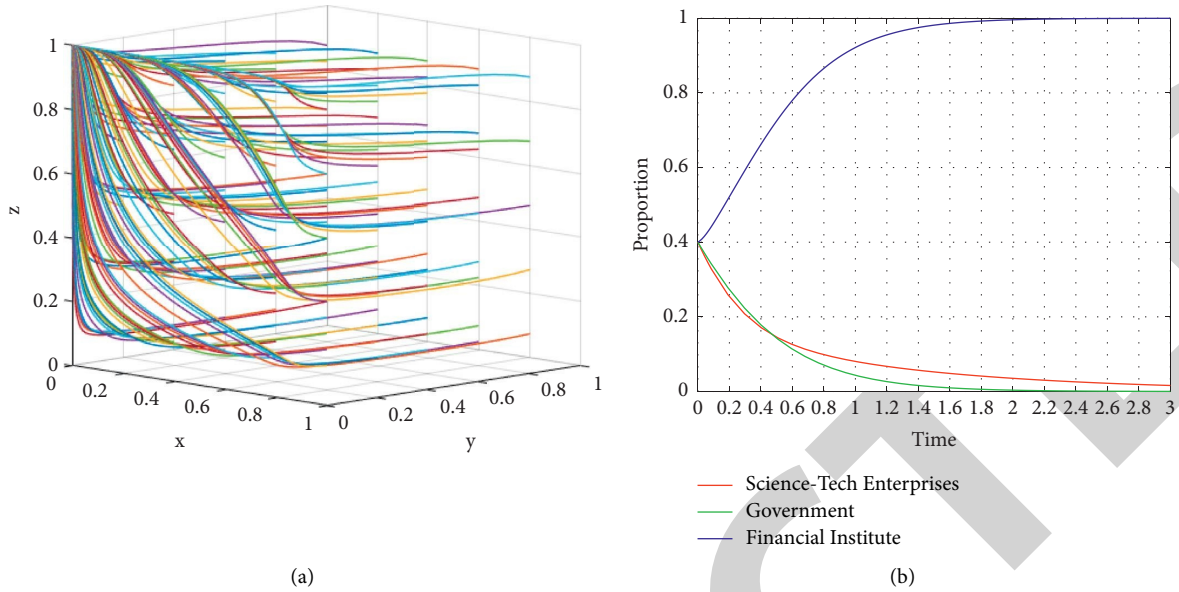


FIGURE 8: (a) Three-dimensional situation of evolutionary trajectory of A (0, 0, 1). (b) Two-dimensional situation of evolutionary trajectory of A (0, 0, 1).

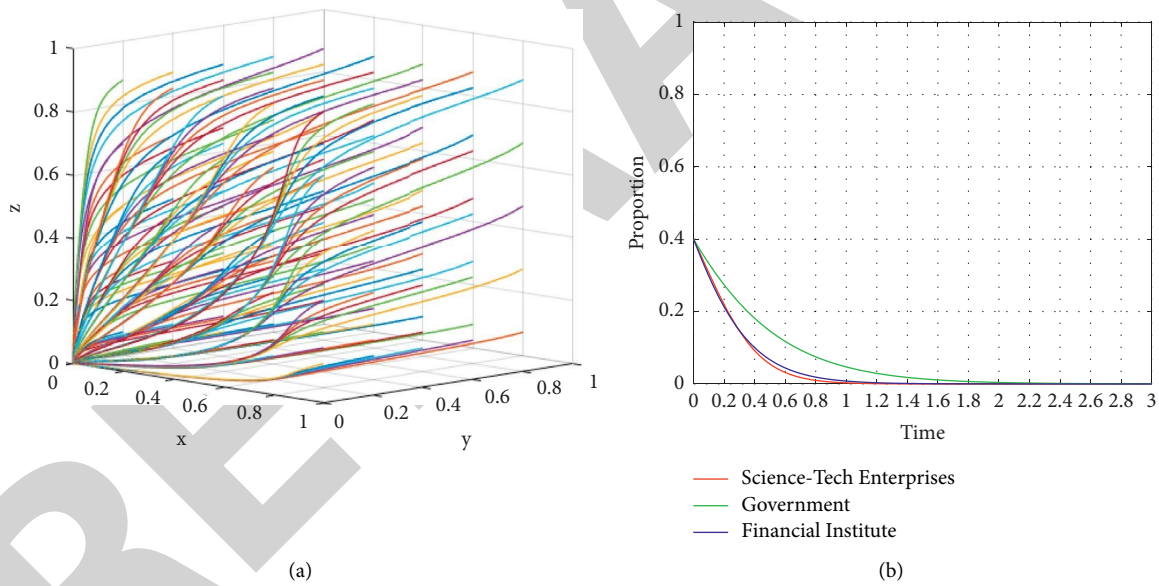


FIGURE 9: (a) Three-dimensional situation of evolutionary trajectory of O (0, 0, 0). (b) Two-dimensional situation of evolutionary trajectory of O (0, 0, 0).

in Figure 5(a). When the initial probabilities of all three parties are 0.4, the evolutionary trajectory is displayed in Figure 5(b).

- (5) Assume  $F_Q = 100$ ,  $F_L = 95$ ,  $\alpha_Q = 0.3$ ,  $\alpha_L = 0.2$ ,  $\beta_Q = 1.05$ ,  $\beta_L = 1.7$ ,  $S_Q = 10$ ,  $S_L = 8$ ,  $k = 1.2$ ,  $C_G = 1$ ,  $g_Q = 1.2$ ,  $g_L = 1.1$ ,  $\delta = 0.3$ ,  $r_S = 0.05$ , and  $r_N = 0.08$ . The evolutionary trajectory of C (0, 1, 1) is featured in Figure 6(a). When the initial probabilities of all three parties are 0.4, the evolutionary trajectory is displayed in Figure 6(b).

- (6) Assume  $F_Q = 100$ ,  $F_L = 95$ ,  $\alpha_Q = 0.3$ ,  $\alpha_L = 0.2$ ,  $\beta_Q = 1.05$ ,  $\beta_L = 1.7$ ,  $S_Q = 10$ ,  $S_L = 8$ ,  $k = 1.2$ ,  $C_G = 1$ ,  $g_Q = 1.2$ ,  $g_L = 1.1$ ,  $\delta = 0.1$ ,  $r_S = 0.06$ , and  $r_N = 0.08$ . The evolutionary trajectory of B (0, 1, 0) is featured in Figure 7(a). When the initial probabilities of all three parties are 0.4, the evolutionary trajectory is displayed in Figure 7(b).

- (7) Assume  $F_Q = 100$ ,  $F_L = 95$ ,  $\alpha_Q = 0.3$ ,  $\alpha_L = 0.2$ ,  $\beta_Q = 1.05$ ,  $\beta_L = 1.7$ ,  $S_Q = 10$ ,  $S_L = 8$ ,  $k = 1.1$ ,  $C_G = 1$ ,  $g_Q = 1.1$ ,  $g_L = 1.05$ ,  $\delta = 0.3$ ,  $r_S = 0.06$ , and  $r_N = 0.07$ .

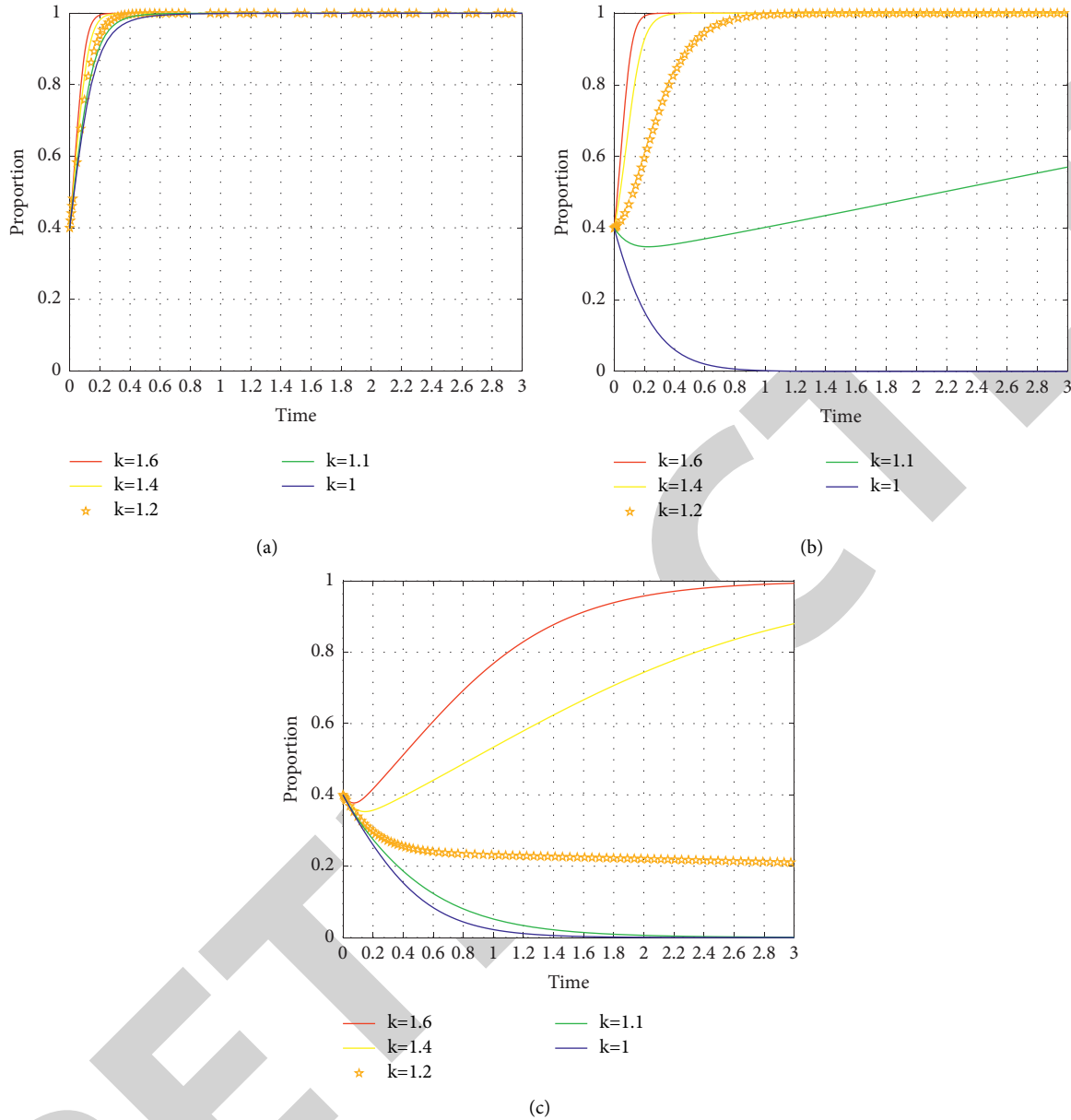


FIGURE 10: Impact of the reputation multiplier by government subsidies “ $k$ ” for (a) science-tech enterprises; (b) for government; (c) for financial institutes.

The evolutionary trajectory of A (0, 0, 1) is featured in Figure 8(a). When the initial probabilities of all three parties are 0.4, the evolutionary trajectory is displayed in Figure 8(b).

- (8) Assume  $F_Q = 100$ ,  $F_L = 95$ ,  $\alpha_Q = 0.3$ ,  $\alpha_L = 0.2$ ,  $\beta_Q = 1.05$ ,  $\beta_L = 1.7$ ,  $S_Q = 10$ ,  $S_L = 8$ ,  $k = 1.1$ ,  $C_G = 1$ ,  $g_Q = 1.1$ ,  $g_L = 1.05$ ,  $\delta = 0.1$ ,  $r_S = 0.06$ , and  $r_N = 0.08$ . The evolutionary trajectory of O (0, 0, 0) is featured in Figure 9(a). When the initial probabilities of all three parties are 0.4, the evolutionary trajectory is displayed in Figure 9(b).

## 5.2. Single-Factor Sensitivity Analysis

### 5.2.1. Impact of Reputation Multiplier by Government Subsidies.

To evaluate the impact of reputation multiplier “ $k$ ” by government subsidies on the evolutionary outcomes and trajectories of the innovation tripartite ESS, numerical simulations were conducted in this study. Set the initial parameters as follows. F (1, 1, 0):  $F_Q = 100$ ,  $F_L = 80$ ,  $\alpha_Q = 0.3$ ,  $\alpha_L = 0.2$ ,  $\beta_Q = 1.5$ ,  $\beta_L = 1.2$ ,  $S_Q = 10$ ,  $S_L = 8$ ,  $k = 1.2$ ,  $C_G = 1$ ,  $g_Q = 1.2$ ,  $g_L = 1.1$ ,  $\delta = 0.1$ ,  $r_S = 0.06$ , and  $r_N = 0.08$ .

Let  $k = 1.6, 1.4, 1.2, 1.1, 1$  and keep other parameters constant; the evolutionary outcomes and trajectories of tripartite ESS are as shown in Figures 10(a)–10(c). This indicates that the size of the reputation multiplier from government subsidies does not affect the R&D intensity of



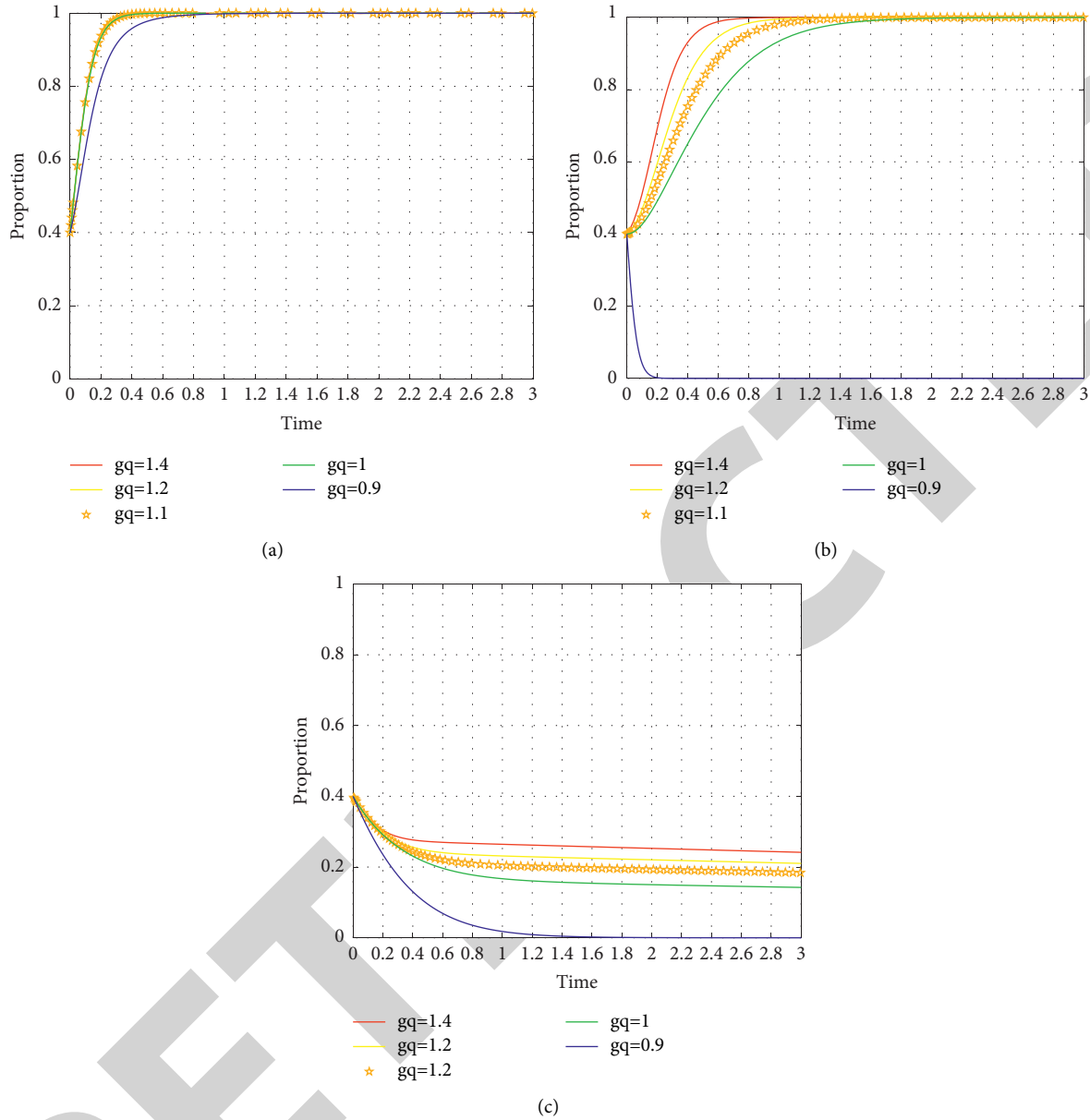


FIGURE 11: Impact of economic and social externality multiplier by innovation of enterprises with high R&D intensity “ $g_Q$ ” (a)  $g_Q$  for science-tech enterprises; (b) for government; (c) for financial institutes.

science-tech enterprises. This means that enterprises will choose to invest in high R&D regardless of whether government subsidies send a positive signal to the market. As the reputation multiplier from government financial subsidies increases for enterprises, the government is more willing to provide subsidies to enterprises. However, for financial institutes, the reputation multiplier from government subsidies affects the investment decisions of financial institutes. As the reputation multiplier decreases, financial institutes will be more inclined to invest in debt and will accelerate the rate of evolution.

**5.2.2. Impact of Economic and Social Externality Multiplier by Innovation.** To evaluate the impact of economic and

social externality multiplier “ $g_Q$ ” by innovation of enterprises with high R&D intensity on the evolutionary outcomes and trajectories of the innovation tripartite ESS, numerical simulations were conducted in this study. Set the initial parameters as follows.  $F(1, 1, 0)$ :  $F_Q = 100$ ,  $F_L = 80$ ,  $\alpha_Q = 0.3$ ,  $\alpha_L = 0.2$ ,  $\beta_Q = 1.5$ ,  $\beta_L = 1.2$ ,  $S_Q = 10$ ,  $S_L = 8$ ,  $k = 1.2$ ,  $C_G = 1$ ,  $g_Q = 1.2$ ,  $g_L = 1.1$ ,  $\delta = 0.1$ ,  $r_S = 0.06$ , and  $r_N = 0.08$ .

Let  $g_Q = 1.4, 1.2, 1.1, 1, 0.9$  and keep other parameters constant; the evolutionary outcomes and trajectories of tripartite ESS are as shown in Figures 11(a)–11(c). This indicates that, as the economic and social externalities from innovation by high R&D intensity enterprises increase, the government is more willing to subsidize enterprises, and enterprises are more willing to choose high R&D intensity. In contrast, a negative economic and social externality from

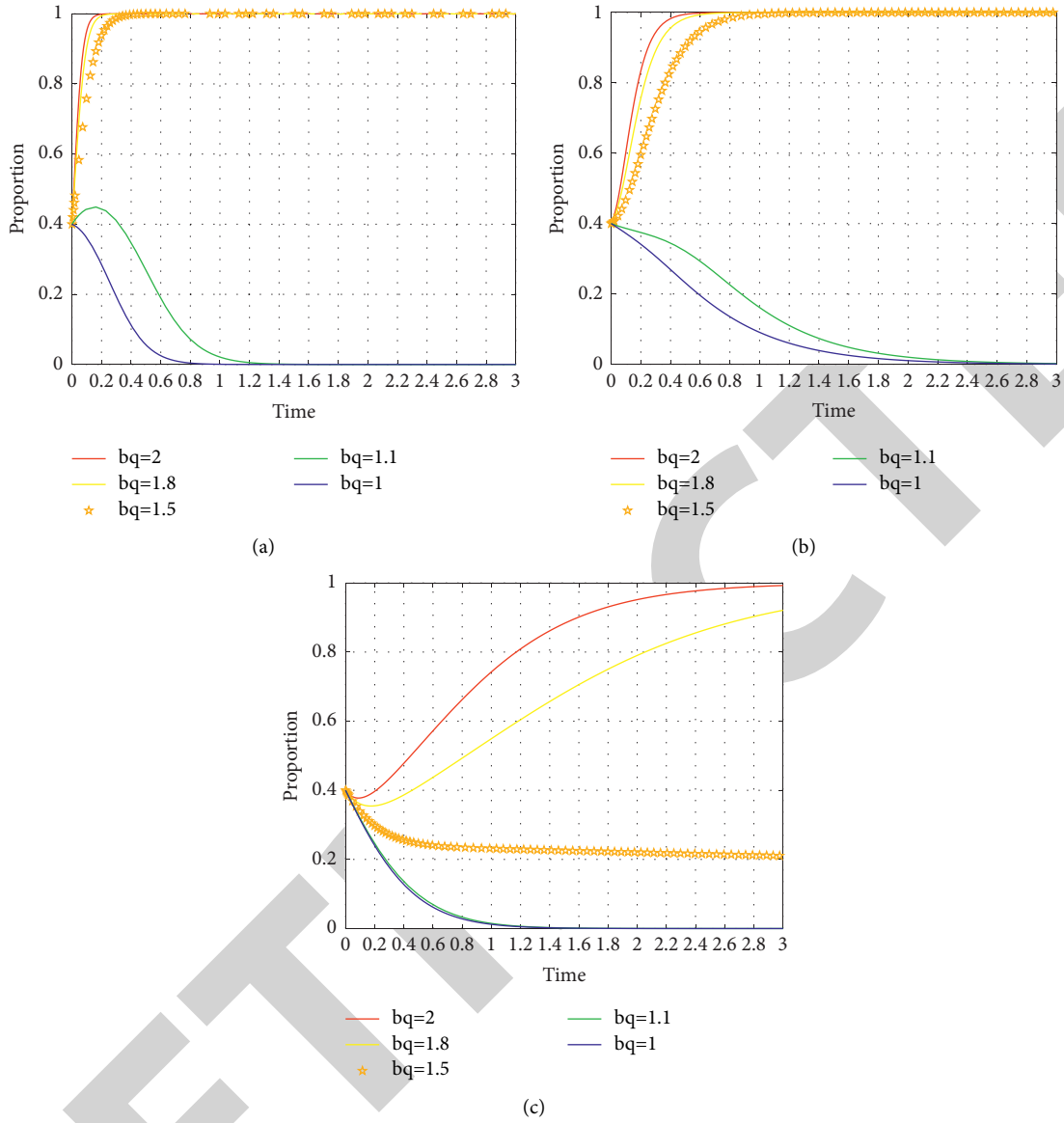


FIGURE 12: Impact of revenue per unit R&D expense for enterprises with high R&D intensity “ $\beta_Q$ ” (a)  $\beta_Q$  for Science-Tech Enterprises; (b) for government; (c) for financial institutes.

a firm’s innovation, i.e., a multiplier less than 1, increases the probability that the government will choose not to subsidize and that it will accelerate the rate of evolution. However, regardless of whether the socio-economic externality is positive or negative, financial institutes tend to choose debt investment to protect against risk.

**5.2.3. Impact of Revenue per Unit R&D Expense for Enterprises with High R&D Intensity.** To evaluate the impact of revenue per unit R&D expense for enterprises with high R&D intensity “ $\beta_Q$ ” on the evolutionary outcomes and trajectories of the innovation tripartite ESS, numerical simulations were conducted in this study. Set the initial parameters as follows.  $F(1, 1, 0)$ :  $F_Q = 100$ ,  $F_L = 80$ ,  $\alpha_Q = 0.3$ ,

$\alpha_L = 0.2$ ,  $\beta_Q = 1.5$ ,  $\beta_L = 1.2$ ,  $S_Q = 10$ ,  $S_L = 8$ ,  $k = 1.2$ ,  $C_G = 1$ ,  $g_Q = 1.2$ ,  $g_L = 1.1$ ,  $\delta = 0.1$ ,  $r_S = 0.06$ , and  $r_N = 0.08$ .

Let  $\beta_Q = 2, 1.8, 1.5, 1.1, 1$  and keep other parameters constant; the evolutionary outcomes and trajectories of tripartite ESS are as shown in Figures 12(a)–12(c). This indicates that science-tech enterprises are more inclined to high R&D intensity as the revenue per unit of R&D expense for enterprises with high R&D intensity, and vice versa. The probability of government subsidies also increases with the revenue per unit of R&D expense for enterprises with high R&D intensity increases. Financial institutes increase the probability of debt investment as the revenue per unit of R&D expense for enterprises with high R&D intensity decreases.

**5.2.4. Impact of Equity Investment Earnings as % of Revenue.**

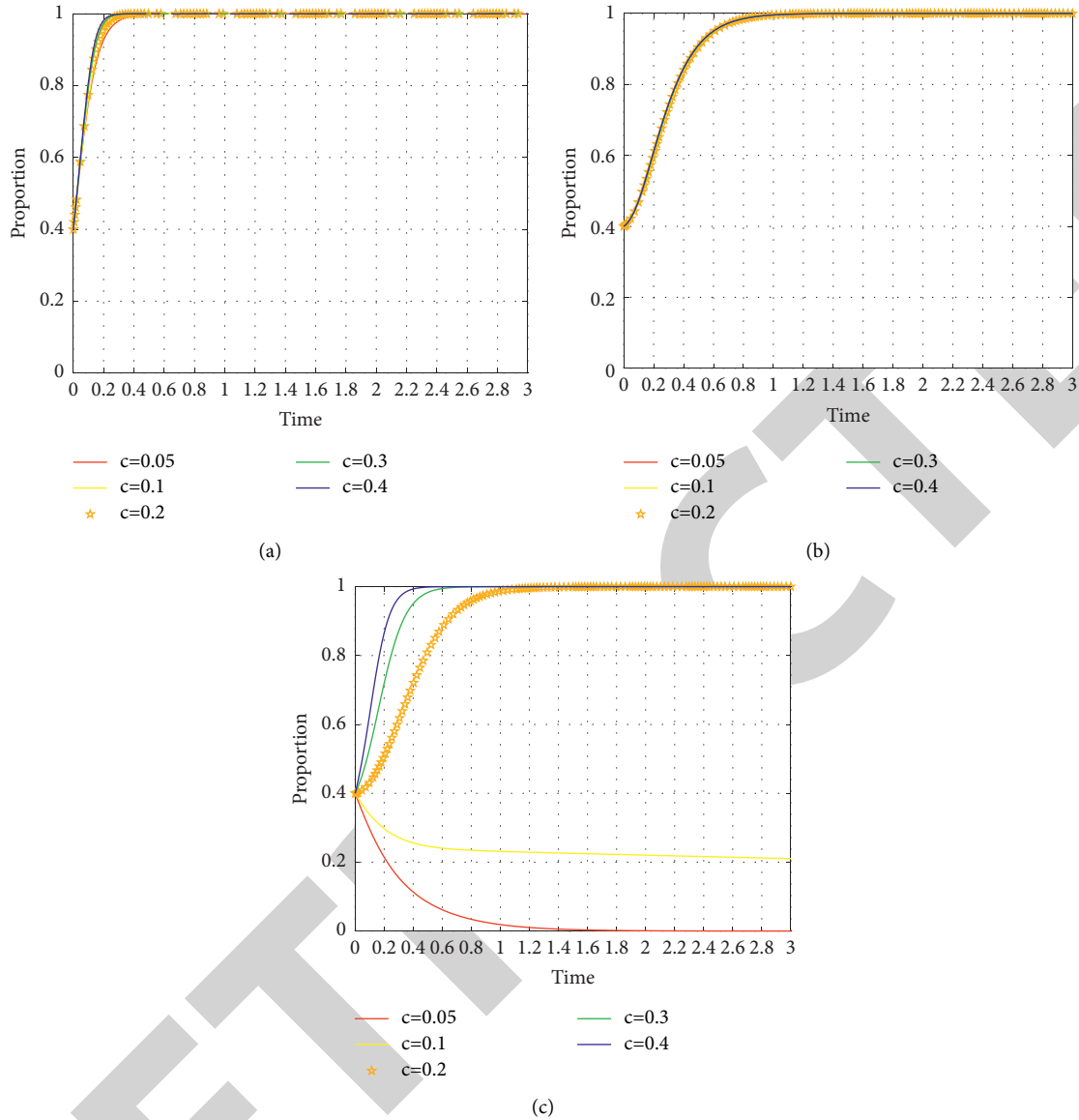


FIGURE 13: Impact of equity investment earnings as % of revenue “ $\delta$ ” (a)  $\delta$  for science-tech enterprises; (b) for government; (c) for financial institutes.

To evaluate the impact of equity investment earnings as % of revenue “ $\delta$ ” for science-tech enterprises “ $\delta$ ” on the evolutionary outcomes and trajectories of the innovation tripartite ESS, numerical simulations were conducted in this study. Set the initial parameters as follows.  $F(1, 1, 0)$ :  $F_Q = 100$ ,  $F_L = 80$ ,  $\alpha_Q = 0.3$ ,  $\alpha_L = 0.2$ ,  $\beta_Q = 1.5$ ,  $\beta_L = 1.2$ ,  $S_Q = 10$ ,  $S_L = 8$ ,  $k = 1.2$ ,  $C_G = 1$ ,  $g_Q = 1.2$ ,  $g_L = 1.1$ ,  $\delta = 0.1$ ,  $r_S = 0.06$ , and  $r_N = 0.08$ .

Let  $\delta = 0.05, 0.1, 0.2, 0.3, 0.4$  and keep other parameters constant; the evolutionary outcomes and trajectories of tripartite ESS are as shown in Figures 13(a)–13(c). This indicates that the probability of choice by firms and government is not affected by equity investment earnings as % of revenue. Regardless of the return on equity investment,

science-tech enterprises will tend to choose high R&D intensity, while the government will tend to choose subsidies. However, as the rate of return on equity investment increases, financial institutes will increase the probability of making equity investments in science-tech enterprises.

## 6. Conclusion

This study focuses on the tripartite decision-making mechanism between government, financial institutes, and technology companies on subsidies, investment patterns, and R&D intensity. The following conclusions are drawn from the study in this paper.

First, enterprises can attract government subsidy by boosting economic and social externalities. The higher density of R & D investment in the science-tech enterprises is, the higher positive external economy and benefits to the government is. As the government pays attention to broad economic and social targets, enterprises can choose the R&D projects with high positive economic and social externalities to provide incentive to the government to provide subsidy.

Second, government subsidy policy can influence the choices of enterprises and financial institutes. The government subsidies to enterprises with high R&D intensity will improve the innovation output by the innovation investment and reputation multiplier. So, the government can implement subsidy policy to attract enterprises to make high-intensive R&D activity and financial institutes to make equity investment. For small high-tech startups, government subsidy can improve their reputation in the market, which can help them to raise fund in the market and find good partners and increase the R&D output efficiency. It would be a better approach for the government to improve the market credible evaluation and subsidy recognition system for high-tech companies as a way to increase their reputation multiplier. In reality, government subsidies can take many forms and tax credit, for R&D expense is a popular policy.

Third, financial institutes' strategy choice is influenced by investment yield difference and can influence enterprises' R&D intensity choices. The policymakers can boost equity investment and R&D activity by developing capital market and optimizing subsidy policy. Supporting enterprises with high R&D intensity to go listing can attract financial institutes to make equity investment and enterprises to choose high R&D intensity strategy.

The above research results provide a good reference for policymakers to develop strategies for innovation. However, some limitations of this study still exist: this study only considers the effect of government subsidy behavior on the choice of investment patterns of financial institutions and the R&D intensity of firms, without further examining the effect of other incentive approaches. At the same time, research institutes (including universities and other institutes) need to be included in the analytical framework for exploration. Thus, multiparty decisions under multiple government incentive discretionary decisions will be the next step of research.

### Data Availability

All the data used to support the findings of this study are available from the corresponding author upon request.

### Conflicts of Interest

The authors declare that they have no conflicts of interest.

### Acknowledgments

This study was financially funded by the project of Beijing Philosophy and Social Science Foundation of the year 2016 (16YJC045) "Research on the influence mechanism of

Beijing Science and technology finance network collaboration on enterprise innovation performance."

### References

- [1] R. T. Frambach and N. Schillewaert, "Organizational innovation adoption: a multi-level framework of determinants and opportunities for future research," *Journal of Business Research*, vol. 55, no. 2, pp. 163–176, 2002.
- [2] J. R. Brown, S. M. Fazzari, and B. C. Petersen, "Financing innovation and growth: cash flow, external equity, and the 1990s R&D boom," *The Journal of Finance*, vol. 64, no. 1, pp. 151–185, 2009.
- [3] B.-H. Hall, M.- Paternò-Castello Pietro, M. Sandro, and V. Antonio, "Financing constraints, R&D investments and innovative performances: new empirical evidence at the firm level for Europe," *Economics of Innovation and New Technology*, vol. 25, 2016.
- [4] B. H. Hall, "The financing of research and development," *National Bureau of Economic*, vol. 18, no. 1, pp. 35–51, 2002.
- [5] A. Fu-Xin, B. Jie-Zhang, and C. Wenping-Zheng, "Does credit market impede innovation? Based on the banking structure analysis," *International Review of Economics & Finance*, vol. 52, Article ID 52268, 2017.
- [6] F. Helmut, K. Karsten, and U. Katrin, "The interdependence of R&D activity and debt financing of young firms," *Journal of Small Business Management*, vol. 53, no. Suppl S1, pp. 251–277, 2015.
- [7] A. Zh, B. Gl, and A. Zl, "Loaning scale and government subsidy for promoting green innovation," *Technological Forecasting and Social Change*, vol. 144, Article ID 144148, 2019.
- [8] F. Yu, Q. Chao, and S.-O. Business, "Threshold effects of government R&D subsidies on R&D investment of technological SMEs," *Science Research Management*, vol. 38, 2017.
- [9] Y. Zhou, "Research on the performance, causes and governance of structural characteristics of China's corporate innovation financing constraints," *Management World*, vol. 5, no. 4, p. 2, 2017.
- [10] T. H. Sabrina, "Financing innovation: evidence from R&D grants," *The American Economic Review*, vol. 107, 2017.
- [11] T. Takalo and T. Tanayama, "Adverse selection and financing of innovation: is there a need for R&D subsidies?" *The Journal of Technology Transfer*, vol. 35, no. 1, pp. 16–41, 2010.
- [12] R. Kleer, "Government R&D subsidies as a signal for private investors," *Research Policy*, vol. 39, 2010.
- [13] N. Yoshino, Emeritus, and F. Taghizadeh-Hesary, "Optimal Credit Guarantee Ratio for Small and Medium-Sized Enterprises' Financing: Evidence from Asia," *Economic Analysis and Policy*, vol. 62, 2018.
- [14] M. Wang and S.-O. Economics, "Research on the risk sharing ratio of the guarantee loans in Micro, Small and medium enterprises—based on the two party cooperation game model with government subsidies," *Journal of Business Economics*, vol. 37, 2017.
- [15] F. M. Menezes and J. Pereira, "Emissions abatement R&D: dynamic competition in supply schedules," *Journal of Public Economic Theory*, vol. 19, no. 4, pp. 841–859, 2017.
- [16] H. Hottenrott and B. Peters, "Innovative capability and financing constraints for innovation: more money, more innovation?" *The Review of Economics and Statistics*, vol. 94, 2012.
- [17] F. Savignac, "Impact OF financial constraints ON innovation: what can be learned from a direct measure?" *Economics of*

- Innovation and New Technology*, vol. 17, no. 5, pp. 553–569, 2008.
- [18] M. Almus and D. Czarnitzki, “The effects of public R&D subsidies on firms’ innovation activities,” *Journal of Business & Economic Statistics*, vol. 21, no. 2, pp. 226–236, 2003.
- [19] D. Emmanuel, “Are R&D subsidies a substitute or a complement to privately funded R&D? Evidence from France using propensity score methods for non- experimental data,” *Public Economics*, vol. 114, no. 411007, p. 245, 2004.
- [20] D. Czarnitzki and A. Fier, “Do innovation subsidies crowd out private investment? Evidence from the German service sector,” *Applied Economics Quarterly*, vol. 48, no. 2-04, 2002.
- [21] D. Aboal and P. Garda, “Does public financial support to innovation increase innovation and productivity?: an impact evaluation,” *Centro De Investigaciones Económicas Montevideo Uy*, 2013.
- [22] G. Saint-Paul, “Technological choice, financial markets and economic development[J],” *European Economic Review*, vol. 36, no. 4, pp. 763–781, 1990.
- [23] H. Luong, F. Moshirian, L. Nguyen, T. Xuan, and Z. Bohui, “How do foreign institutional investors enhance firm innovation?” *Social Science Electronic Publishing*, 2016.
- [24] S. Kim, H. Lee, and J. Kim, “Divergent effects of external financing on technology innovation activity: Korean evidence,” *Technological Forecasting and Social Change*, vol. 106, no. May, pp. 22–30, 2016.
- [25] Q. Gong, Y. Zhang, and Y. Lin, “Industrial structure, risk characteristics and optimal financial structure[J],” *Economic Research*, vol. 49, no. 04, pp. 4–16, 2014.
- [26] M. P. Feldman and M. R. Kelley, “The ex ante assessment of knowledge spillovers: government R&D policy, economic incentives and private firm behavior,” *Research Policy*, vol. 35, no. 10, pp. 1509–1521, 2006.
- [27] M. Meuleman and W. De Maeseneire, “Do R&D subsidies affect SMEs’ access to external financing?” *Research Policy*, vol. 41, no. 3, pp. 580–591, 2012.
- [28] A. Wu, “The signal effect of Government R&D Subsidies in China: does ownership matter?” *Technological Forecasting and Social Change*, vol. 117, no. APR, pp. 339–345, 2017.
- [29] J. Lerner, “The Government as Venture Capitalist: the long-run impact of the SIBR program,” *Nber Working Papers*, vol. 72, no. 3, pp. 285–318, 1996.
- [30] D. Guellec and B. Van Pottelsberghe De La Potterie, “The impact of public R&D expenditure on business R&D\*,” *Economics of Innovation and New Technology*, vol. 12, no. 3, pp. 225–243, 2003.
- [31] X. Liu, X. Li, and H. Li, “R&D subsidies and business R&D: evidence from high-tech manufacturing firms in Jiangsu,” *China Economic Review*, vol. 41, pp. 411–422, 2016.
- [32] Z. Yang, Y. Shi, and Y. Li, “Analysis of intellectual property cooperation behavior and its simulation under two types of scenarios using evolutionary game theory,” *Computers & Industrial Engineering*, vol. 125, no. NOV, pp. 739–750, 2018.
- [33] Y. Han, G. Chen, and E. Poh, “Effects of informal contracts on innovative cooperation among enterprises in industrial clusters: an evolutionary game analysis,” *Discrete Dynamics in Nature and Society*, vol. 2018, Article ID 5267357, 10 pages, 2018.
- [34] Bo. Shen, “The influence of endogenous knowledge spillovers on open innovation cooperation modes selection,” *Wireless Personal Communications*, vol. 102, 2018.
- [35] J. Ma, T. Xu, Y. Hong, and Z. Xueli, “Impact research on a nonlinear cold chain evolutionary game under three various contracts,” *International Journal of Bifurcation and Chaos*, vol. 29, no. 5, Article ID 1950058, 2019.
- [36] X. Su, H. Liu, and S. Hou, “Corrigendum to “the trilateral evolutionary game of agri-food quality in farmer-supermarket direct purchase: a simulation approach,”” *Complexity*, vol. 2019, Article ID 5163589, 1 page, 2019.
- [37] X. M. Meng-X-X Zhao, “Evolutionary game decision between manufacturer and remanufacturer in the authorization mode,” *Chinese Journal of Management Science*, vol. 29, no. 2, pp. 129–136, 2021.
- [38] B. P. Chen-Wei, “Cooperative innovation of key common technologies in equipment manufacturing industry based on evolutionary game,” *Enterprise economy*, no. 7, pp. 25–33, 2019.
- [39] P. Vaida, “R&D investment and competitiveness in the baltic states,” *Procedia Social & Behavioral Sciences*, vol. 213, 2015.
- [40] S. Zhang, “The tripartite evolution game of manufacturing enterprise innovation strategy under the participation of the government,” *Journal of North China University of Water Resources and Electric Power (SOCIAL SCIENCE EDITION)*, vol. 38, no. 1, pp. 18–25, 2022.

## Retraction

# Retracted: Feature Sequencing Method of Industrial Control Data Set Based on Multidimensional Evaluation Parameters

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.









The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] X. Liu, X. Kong, X. Zhang et al., "Feature Sequencing Method of Industrial Control Data Set Based on Multidimensional Evaluation Parameters," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 9248267, 10 pages, 2022.

## Research Article

# Feature Sequencing Method of Industrial Control Data Set Based on Multidimensional Evaluation Parameters

Xue-Jun Liu <sup>1</sup>, Xiang-Min Kong <sup>1</sup>, Xiao-Ni Zhang <sup>1</sup>, Hai-Ying Luan,<sup>2</sup> Yong Yan <sup>1</sup>, Yun Sha <sup>1</sup>, Kai-Li Li <sup>1</sup>, Xue-Ying Cao <sup>1</sup>, and Jian-Ping Chen <sup>1</sup>

<sup>1</sup>College of Information Engineering, Beijing Institute of Petrochemical Technology, 19 Qingyuan North Road, Daxing District, Beijing, China

<sup>2</sup>Fluid Drive and Car Equipment Technical Engineering Department, Beijing Research Institute of Automation for Machinery Industry Co., Ltd, 100120 Beijing, China

Correspondence should be addressed to Xue-Jun Liu; [lxj@bipt.edu.cn](mailto:lxj@bipt.edu.cn)

Received 3 February 2022; Revised 15 March 2022; Accepted 31 March 2022; Published 28 April 2022

Academic Editor: Daqing Gong

Copyright © 2022 Xue-Jun Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The industrial control data set has many features and large redundancy, which has a certain impact on the training speed and classification results of the neural network anomaly detection algorithm. However, features are independent of each other, and dimension reduction often increases the false positive rate and false negative rate. The feature sequencing algorithm can reduce this effect. In order to select the appropriate feature sequencing algorithm for different data sets, this paper proposes an adaptive feature sequencing method based on data set evaluation index parameters. Firstly, the evaluation index system is constructed by the basic information of the data set, the mathematical characteristics of the data set, and the association degree of the data set. Then, the selection model is obtained by the decision tree training with the data label and the evaluation index, and the suitable feature sequencing algorithm is selected. Experiments were conducted on 11 data sets, including Batadal data set, CICIDS 2017, and Mississippi data set. The sequenced data sets are classified by ResNet. The accuracy of the sequenced data sets increases by 2.568% on average in 30 generations, and the average time reduction per epoch is 24.143%. Experiments show that this method can effectively select the feature sequencing algorithm with the best comprehensive performance.

## 1. Introduction

With the development of industrial control systems [1] and digital communication technology, the industrial control network needs to face more and more external network access attacks [2–3]. Therefore, more research has been carried out on anomaly detection algorithm [4–8]. However, in the process of digitization, the feature dimensions in industrial control data increase, which increases the complexity of data processing tasks. This leads to the increase in learning cost and memory cost in anomaly detection, which limits the establishment of the learning model [9]. How to reduce the complexity between features and speed up the establishment of the model has become an urgent problem to be solved.

Methods in feature engineering are often used to reduce the complexity of features. Porizka et al. [10] applied the

principal component analysis algorithm to the Laser-Induced Breakdown Spectroscopy to process the detected multivariate signals (characteristic spectrum), but the principal component analysis algorithm will map the characteristic data, and the results are different from the original features. Chen et al. [11] used the ant colony algorithm to select features, eliminate redundant features, and improve the speed of deep reinforcement learning training, but the generalization ability of the model decreases after reducing features.

Feature sequencing methods have been widely used in various engineering projects, such as filtering method [12], Pearson correlation coefficient, Spearman correlation coefficient [13], information entropy [14], Lasso [15], elastic network [16], recursive feature elimination based on SVM [17], Bayesian kernel model [18], and gradient learning [19]. In addition, there are many studies on the application



in specific scenes. Pandeya et al. [20] applied the REF algorithm to the risk feature identification in the banking industry, designated the features of the data according to the value of credit risk classification, evaluated its value through repeated sampling, and gave weight. Then, these weights are used to separate the adjacent data of the same and different credit risks to complete the risk feature sequencing. In the structural damage detection of civil housing, Zhou et al. [21] processed the collected vibration acceleration signal through wavelet packet decomposition and converted it into the initial energy feature set, then eliminated the least important features through the RF-REF algorithm, evaluated the importance of the important sequence of features, and completed the reordering of features. Matej et al. [22] studied the problem of finding biomarkers, used random forest and RReliefF algorithms for feature sequencing, and compared the forward feature addition curve and reverse feature addition curve of the two algorithms in two different data sets, as well as the sequencing stability.

The above references all introduce feature sequencing algorithms, but the effect of the feature sequencing method on different data sets is obviously different. When the consistency between the method and the application scene features is low, there will be problems such as selection redundancy, unable to detect the relationship between all features, and high computational complexity [23, 24]. This paper presents a method to quickly find an appropriate feature reordering algorithm for data sets.

In view of the above problems, the main objectives of this paper could be summarized as follows: (1) Establish the evaluation index system of basic information of the data set. (2) Complete the adaptive feature sequencing method.

The current work can be divided into five sections. Section 1 introduces the necessity of improving the feature sequencing algorithm in an industrial control system. Section 2 describes the methods used in the current work. Then, Section 3 describes the experimental design. Section 4 describes the results, and the discussion and conclusion are described in detail in Section 5.

## 2. Method

In the field of industrial control anomaly detection, the feature sequencing algorithm is a preprocessing algorithm used in neural networks to detect system parameter information, mainly to solve the impact of the correlation uncertainty between dimensions on the neural network algorithm. The algorithm in this paper is mainly divided into an evaluation index system and decision tree model construction, as shown in Figure 1.

In this paper, the experimental data set is input, and the sequenced data set after sequencing is obtained by different feature sequencing algorithms. Then, the parameters of the sequenced data set in the evaluation index system are calculated. Finally, the Gini coefficient is used to construct the decision tree model to complete the selection method and realize the program function. The specific steps are shown in Figure 2.

*2.1. Evaluation Index System.* After investigation, a set of evaluation index systems was established to analyze the characteristics of different data sets from multiple perspectives [25, 26]. In this paper, the evaluation index system is constructed from the basic information of the data set, the mathematical characteristics of the data set, and the degree of association of the data set. The mathematical characteristics include data distribution, data association, and data collinearity. Finally, the evaluation indexes include 8 indexes in 3 categories and 5 subcategories. For different data sets and different feature sequencing algorithms, the evaluation indexes are used to evaluate the data sets. The indicators are shown in Table 1.

### 2.1.1. Evaluation Indexes

(1) *The Number of Dimensions.* The number of dimensions can reflect the complexity of the data set. Generally speaking, the more dimensions of the data set, the more information the data set contains.

(2) *The Number of Categories.* The number of categories is different in data sets. In the case of similar data collective quantity, the more categories the numbers represent, the more the data types contained and the higher the complexity of the data set. The number of classifications will directly affect the algorithm's detection effect of the data set. Therefore, the number of classifications is also used as the evaluation index parameter of the data set.

(3) *Variance.* Variance is a disperse measurement of a random variable or set of data in probability theory and statistics. Large variance indicates high degree of data dispersion, and small variance indicates strong degree of data aggregation.

(4) *The Imbalance Ratio between Categories.* In multi-classification data sets, each category contains different sample sizes. This situation leads to unbalanced distribution among data categories, which can be measured by the imbalance ratio (IR) between categories. The larger the value of IR is, the more unbalanced the category distribution of sample data is, which will easily affect the classification accuracy.

(5) *KL Divergence.* KL divergence is used to calculate the cumulative difference between the information entropy of real events and the information entropy of theoretical fitting. It can be used to measure the distance between two dimensional distributions. When the distribution in two dimensions is the same, KL divergence is zero. When the difference of the two dimensional distributions increases, KL divergence also increases.

(6) *Curve Fitting Degree.* Curve fitting degree (CFD) is also an embodiment of data repetition. The degree of data trend repetition can be measured by CFD. The degree of data redundancy in industrial control system can be calculated by CFD.

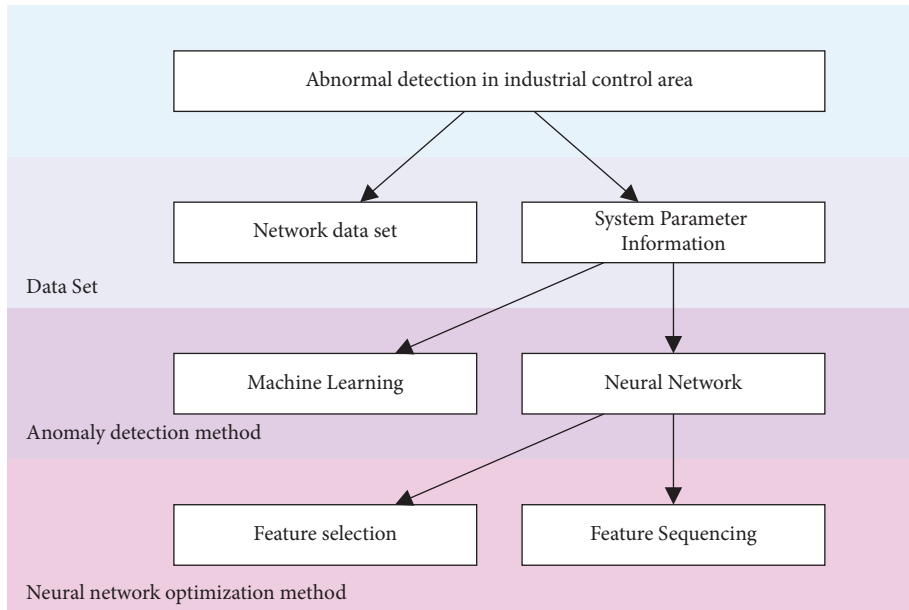


FIGURE 1: Technical subdivision of the industrial control field.

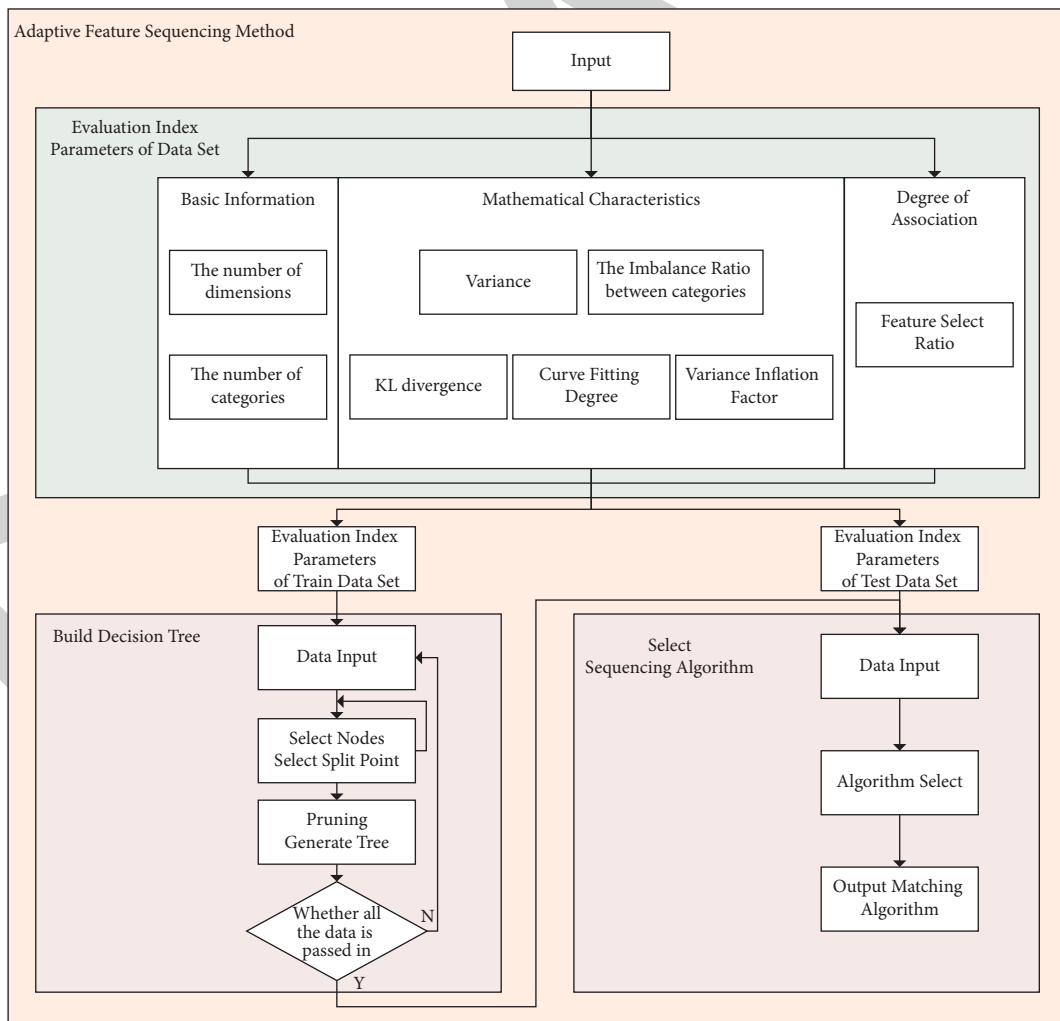


FIGURE 2: Diagram of the main program.

TABLE 1: Evaluation indexes.

Classification	Evaluation index	Abbreviation
Basic information	The number of dimensions	Dimension
	The number of categories	Category
Mathematical characteristics, data distribution	Variance	\
	The imbalance ratio between categories	IR
Mathematical characteristics, data association	KL divergence	KL
	Curve fitting degree	CFD
Mathematical characteristics, data collinearity	Variance inflation factor	VIF
	Degree of association	FS-R

(7) *Variance Inflation Coefficient*. Variance inflation coefficient (VIF) tests the linear correlation between features of the data set. This index parameter can select features with strong independence and increase the interpretability of the model.

(8) *Feature Select Ratio*. Feature selection can calculate the degree of association between dimensions, so the proportion of feature selection is taken as the evaluation index. When the feature selection method is used for feature screening of data sets, the number of retained features of different data sets is different. For example, the Lasso algorithm was used to select features from the Mississippi data set [27] and only 4 features were retained and 22 features were deleted. The same feature selection method was used for feature selection of the CICIDOS 2019 data set, and 34 features were retained and 4 features were deleted. The same feature selection method has a significant difference in the proportion of deleted features in the data set. Therefore, the proportion of retained features by feature selection is directly related to the characteristics of the data set, which can be used to measure the gap between data sets. The calculation formula of feature selection ratio is as follows:

$$R_{FS} = \frac{N_{FS}}{N}, \quad (1)$$

where  $N$  is the number of dimensions of the data set,  $N_{FS}$  is the number of features obtained by feature selection of the data set through Lasso algorithm, and  $R_{FS}$  is the proportion of features selected by feature selection algorithm in all dimensions.

2.1.2. *Calculation of the Evaluation Index System*. According to the feature sequencing method mentioned above, different methods were used to conduct feature sequencing for each data set and then the above-mentioned

evaluation indexes were calculated. The evaluation index parameter result of each data set was calculated as the parameters of this data set. After all data sets were calculated, the evaluation index parameter set was obtained.

In order to verify the above-mentioned evaluation index, five different data sets were selected to calculate the above-mentioned indicators, respectively. Data sets include the Mississippi data set [27], Oil Depot data set, the CICIDOS 2017 [28], the Wine data set [29], and the Csgo data set [30]. The results are shown in Figure 3.

As shown in Figure 3, the above-mentioned evaluation indexes have great differences in different data sets. And, indexes are highly independent of each other in distribution, which can distinguish different data sets. When using decision tree algorithm for classification, each index parameter can be used as a feature to construct nodes and then the parent-child relationship between nodes can be constructed according to the calculation results of evaluation indexes. The selected index parameter can reflect the situation of data sets from different angles and is suitable for decision tree algorithm.

2.1.3. *Label*. The data sets are labeled according to the sequencing algorithm of different features. According to the accuracy and time of the classification algorithm in the sequencing data set, the calculation formula of identification principle is as follows:

$$Acc_{ix} = \text{MAX}(Acc_{i1}, Acc_{i2}, \dots, Acc_{ij}). \quad (2)$$

$Acc_{ij}$  is the accuracy of abnormal detection neural network after the JTH feature sequencing algorithm is adopted for the Ith data set, and  $x$  is the selection method to select the  $x$  method for feature sequencing. When the same data set has more than one of the same highest accuracy (usually 100% accuracy at the same time), we use time to identify. The calculation formula is as follows:

$$\begin{cases} Acc_{ix1} = Acc_{ix2} = \dots = Acc_{ixk} = \text{MAX}(Acc_{i1}, Acc_{i2}, \dots, Acc_{ij}), \\ Time_{ix} = \text{MAX}(Time_{ix1}, Time_{ix2}, \dots, Time_{ixk}). \end{cases} \quad (3)$$

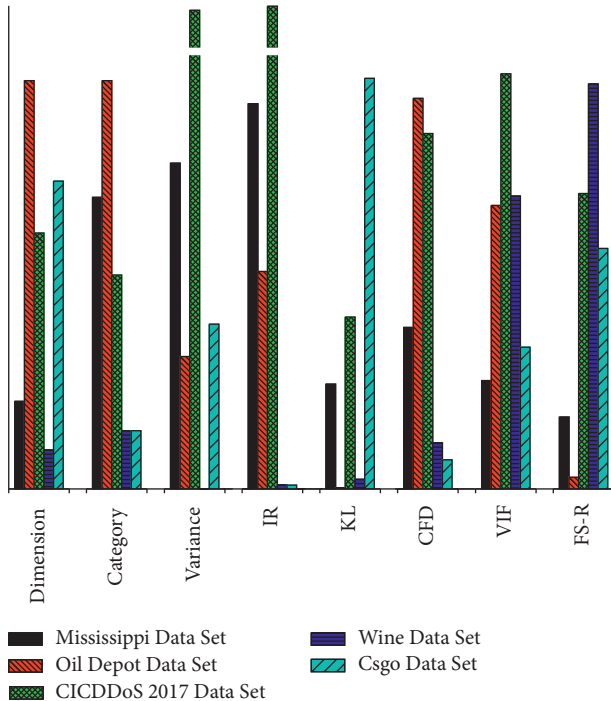


FIGURE 3: Effect pictures of evaluation indexes in different data sets.

$Time_{ixk}$  is the  $K$  TH of the sequencing algorithm with the highest accuracy among the remaining feature sequencing algorithms for the  $I$ th data set and the time of each substitution of neural network anomaly detection. After the selection result is method  $X$ , the data set is labeled as  $X$ .

**2.2. Feature Sequencing Algorithm.** Feature sequencing algorithm can find the dimension with high importance, and it can sequence features according to the importance score to change the distance between features. By this method, we can solve the problem of uncertain correlation between adjacent dimensions of industrial control system data and achieve convergence acceleration effect of anomaly detection algorithm through feature sequencing algorithm. In the previous experiment, Lasso regularization feature selection algorithm was used for feature sequencing, but the algorithm has different effects on different data sets. Therefore, this paper selects several common feature processing methods as feature sequencing algorithm for experiment.

Common sequencing methods include feature selection method, regularization method, random forest method, and top-level selection method. After the investigation, this paper selected a total of 4 categories and 7 algorithms, including Pearson correlation coefficient [13], linear regression, L1 regularization [31], L2 regularization [32], random forest [33], stability selection top-level selection algorithm [34], and recursive feature elimination top-level selection algorithm [35]. The experiment of selecting suitable feature sequence algorithm is carried out. The selection algorithm is shown in Table 2.

TABLE 2: List of feature sequencing algorithms.

Category	Algorithm	Abbreviation
Feature selection method	Pearson correlation coefficient	Pearson
	Linear regression	Linear
Regularization method	Lasso regularization	Lasso
	Ridge regularization	Ridge
Random forest method	Random forest	RF
Top-level selection method	Stability selection	SS
	Recursive feature elimination	RFE

**2.3. Decision Tree Model Construction.** Decision trees [36–38] can use complex nonlinear models to fit the data and change the measurement of impurity for regression analysis. Similar to the linear regression model, corresponding loss function is used and decision tree is used for regression to measure impurity [39]. The decision tree is constructed by taking evaluation indexes of each data set as nodes, and the most suitable feature selection algorithm is selected for different data sets after the evaluation index parameter sets and labels are passed in. The specific steps are as follows.

(1) *Feature Splitting Node.* All features are traversed, and the change value of information entropy before and after dividing the data set is calculated by  $H(X)$ . Then, the feature with the largest change of information entropy is selected as the basis for dividing the data set, that is, the feature with the largest information gain is selected as the split node.

$$H(X) = - \sum_{x \in X} p(x) \log p(x). \quad (4)$$

Here,  $p(x)$  represents the probability of the occurrence of element  $x$  in the dimension. When the probability  $p(x)$  is closer to 0 or 1, the value of information entropy is smaller. When the probability value is 1, the information entropy is 0 and the data category is single. In feature selection, the feature with the maximum information gain is selected, which physically makes the data transform in a single direction as far as possible. Information gain is a measure of the degree to which data become more sequenced.

(2) *Decision Tree Construction.* Firstly, the information entropy before data set partition is calculated. Secondly, the information entropy after dividing the data set according to each feature is calculated, and the feature with the largest information gain is selected as the data partition node to divide the data. Finally, all sub-data sets after partition are recursively processed, and the above steps are repeated from the features that have not been selected to select the optimal data partition feature to partition the molecular data set.

Recursion generally ends under two conditions: all features have been used or the information entropy gain after partition is small enough, that is, as many divided numbers as possible belong to the same category.

(3) *Decision Tree Pruning*. Due to the influence of noise and other factors, the values of some features of the sample do not match the categories of the sample itself and some branches and leaves of the decision tree generated based on these data will produce some errors. Especially in the decision tree near the end of the branches and leaves, due to fewer samples, the interference of irrelevant factors becomes more prominent. The resulting decision tree may be over-fitting, so the classification speed and accuracy of the whole decision tree can be improved by deleting unreliable branches by pruning.

After the establishment of the decision tree, it only needs to input the result data of the evaluation index parameter to select a matching feature sequencing algorithm, thus completing the design of the adaptive algorithm.

### 3. Experiment Details

The experiment environment is as follows:

Operating System Windows Server 2016 Datacenter  
 CPU Intel(R) Xeon(R) CPU E5-2650 v4 @ 2.20 GHz  
 GPU NVIDIA GeForce GTX 1080 Ti  
 Runtime Environment Python 3.8  
 Pytorch 1.7.0

*3.1. Data Set*. There were 11 data sets used in experiment, including Batadal data set [40], CICIDS 2017 [28], Mississippi data set [27], Oil Depot data set (self-build data set), Csgo data set [30], Mail data set [41], Water Quality data set [42], Wine data set [29], Mobile Phone Price data set [43], Mnist in Csv data set [44], and Music Genre data set [45].

Batadal data set is part of the Singapore water plant data set, which is established by Itrust, the network security research center of the university of science and technology design of Singapore. The data set is a full physical medium-sized water supply network, c-town water system. The system includes 1 independent reservoir, 5 valves, 5 pump stations, 11 pumps, 7 storage tanks, 388 interfaces, and 423 pipelines. From the aspects of safe water treatment and water supply system, the actual operation state of the system and the real data after the attack are sorted and recorded to form multiple public data sets for competition.

CICIDS 2017, published by the Canadian Institute for Cybersecurity (CIC) at the University of New Brunswick in Fredericton, is one of the several research public data sets for cybersecurity science research. It is an intrusion detection evaluation data set with positive and negative samples.

The Mississippi data set was published by Mississippi State University. In order to study the network traffic of the SCADA system under normal and attacked conditions, Mississippi State University constructed a set of SCADA systems based on all physical objects in 2014 and build a set of standardized data set. The data set includes network flow, process control, and process measurement characteristics of 28 attacks against two laboratory industrial control systems using the MODBUS application layer protocol. The natural gas tank data set is the underlying

business data set containing the attack. The attack types include reconnaissance attacks, response injection attacks, command injection attacks, and Denial-of-Service (DoS) attacks.

In addition, the Oil Depot data set is a data set provided by the cooperative unit, including 132,000 pieces of data and 126 feature dimensions. The positive and negative samples are divided into 11 classes. The other 7 data sets are public data sets released by the KAGGLE platform, and all data sets are in CSV format.

Some basic information of the data set is shown in Table 3.

*3.2. Experimental Procedure*. The experimental steps are as follows:

- (1) 11 data sets were divided into training sets and testing sets, among which 8 training data sets were used for model training and 3 test data sets were used for model testing
- (2) All data set evaluation indexes were calculated, and the specific indicators are shown in Section 2.1.1
- (3) Feature sequencing algorithm is used to preprocess 8 training data sets, and each feature sequencing algorithm generates a sequenced data set according to the original data set
- (4) Sequenced data sets are classified using anomaly detection algorithm to obtain accuracy and running time
- (5) The indicators of the training data set in Step 2 are taken as features, each data set is labeled according to the results of Step 4 as reference, they are input into the decision tree for training, and the selection model is obtained
- (6) The test data and indicators in Step 2 are input as features into the decision tree selection model to obtain the selection results

### 4. Result Analysis

The experimental data sets were divided into training set and testing set. The training set include Batadal data set, Oil Depot data set, Csgo data set, Mail data set, Water Quality data set, Mobile Phone Price data set, Mnist in Csv data set, and Music Genre data set. The testing set includes CICIDS 2017, Mississippi data set, and Wine data set. According to the calculation results of evaluation indexes, the training set is selected for decision tree generation and the feature sequencing algorithm is selected by the generation model.

ResNet anomaly detection algorithm was adopted in the experiment, and most of the final accuracy reached 100% in the processing of the above-mentioned data sets. In order to facilitate comparison, the stable results of iteration 30 generations were used in the experiment and the calculation accuracy and iteration speed were adopted. The results are shown in Table 4.

Table 4 shows the classification results of data sets in the training set by the ResNet algorithm. Accuracy refers to the accuracy of the result, and time refers to the classification



TABLE 3: Data set information.

Name	Dimension	Category
Batadal data set	124	2
CICIDS 2017	79	6
Mississippi data set	27	8
Oil Depot data set	126	11
Csgo data set	95	2
Mail data set	3000	2
Water Quality data set	10	2
Wine data set	12	2
Mobile Phone Price data set	21	4
Mnist in Csv data set	784	10
Music Genre data set	26	10

TABLE 4: Classification results of training data sets in ResNet.

		Pearson	Linear	Lasso	Ridge	RF	SS	RFE
Batadal data set	Accuracy	100.00%	95.833%	95.833%	97.917%	97.917%	93.750%	93.750%
	Time	<b>7.622</b>	7.621	7.625	7.618	7.619	7.622	7.621
Oil Depot data set	Accuracy	<b>94.989%</b>	91.770%	92.060%	91.838%	92.815%	92.077%	92.614%
	Time	<b>3.658</b>	3.703	3.682	3.713	3.693	3.685	3.686
Csgo data set	Accuracy	<b>99.219%</b>	97.344%	98.125%	98.281%	94.375%	97.500%	97.188%
	Time	<b>3.357</b>	3.367	3.365	3.335	3.335	3.364	3.361
Mail data set	Accuracy	100.00%	100.00%	100.00%	100.00%	100.00%	<b>100.00%</b>	100.00%
	Time	0.527	0.527	0.526	0.526	0.527	<b>0.526</b>	0.527
Water Quality data set	Accuracy	94.375%	100.00%	<b>100.00%</b>	100.00%	100.00%	100.00%	100.00%
	Time	2.296	0.068	<b>0.068</b>	0.069	0.069	0.068	0.068
Mobile Phone Price data set	Accuracy	98.854%	98.750%	<b>99.063%</b>	97.396%	98.438%	97.708%	97.917%
	Time	0.654	0.658	<b>0.658</b>	0.660	0.660	0.661	0.654
Mnist in Csv data set	Accuracy	<b>100.00%</b>	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
	Time	<b>1.498</b>	1.504	1.505	1.508	1.504	1.508	1.500
Music Genre data set	Accuracy	<b>99.566%</b>	98.772%	93.149%	97.264%	96.288%	96.479%	95.851%
	Time	<b>0.653</b>	0.663	0.660	0.659	0.660	0.657	0.661

TABLE 5: Classification results of testing data sets in ResNet.

		Pearson	Linear	Lasso	Ridge	RF	SS	RFE
CICIDS 2017	Accuracy	100.00%	99.349%	98.891%	99.193%	98.734%	99.427%	99.359%
	Time	1.035	1.036	1.033	1.032	1.033	1.035	1.034
Mississippi data set	Accuracy	<b>98.532%</b>	92.617%	91.412%	92.227%	92.656%	92.148%	92.026%
	Time	2.484	2.485	2.470	2.551	2.515	2.481	2.482
Wine data set	Accuracy	100.00%	100.00%	100.00%	100.00%	100.00%	<b>100.00%</b>	100.00%
	Time	0.037	0.037	0.037	0.037	0.037	<b>0.037</b>	0.037

time. The bold part is the optimal item label determined according to the identification principle. After confirming the evaluation index parameters and labeling of the training set, the training set was input into the decision tree to complete the construction of the decision tree. Then, the testing set evaluation index parameter data were input to obtain the decision tree selection result of the testing set.

Table 5 shows the classification results of data sets in the testing set by the ResNet algorithm. The bold part is the decision result of the decision tree model. The comparison between the selected result and the results of other methods shows that the feature sequencing algorithm selected by the model on the three testing sets is optimal or better.

After the adaptive algorithm selects the optimal feature sequencing algorithm for sequencing, the results of the anomaly detection algorithm are compared with those before sequencing, as shown in Table 6.

From the bold part in Table 6, it can be seen that the results of 11 data sets after sequencing are generally better than those before and after sequencing except Csgo data set. The comparison of the two indicators is shown in Figure 4.

Figure 4 shows the comparison of accuracy and time before and after the sequencing algorithm for each data set. On the left is the accuracy comparison chart. It can be seen that the accuracy of Csgo data sets after sequencing is slightly lower than before, and the accuracy of other data sets is

TABLE 6: Comparison of data sets before and after feature sequencing algorithm.

Data sets		Before sequencing	After sequencing
Batadal data set	Accuracy	95.833%	100.00%
	Time	7.630	<b>7.622</b>
CICIDS 2017	Accuracy	100.00%	<b>100.00%</b>
	Time	1.688	<b>1.035</b>
Mississippi data set	Accuracy	94.643%	<b>98.532%</b>
	Time	4.273	<b>2.484</b>
Oil Depot data set	Accuracy	89.286%	<b>94.989%</b>
	Time	5.721	<b>3.658</b>
Csgo data set	Accuracy	<b>100.00%</b>	99.219%
	Time	<b>5.445</b>	3.357
Mail data set	Accuracy	100.00%	<b>100.00%</b>
	Time	0.577	<b>0.526</b>
Water Quality data set	Accuracy	100.00%	<b>100.00%</b>
	Time	0.101	<b>0.068</b>
Wine data set	Accuracy	100.00%	<b>100.00%</b>
	Time	0.051	<b>0.037</b>
Mobile Phone Price data set	Accuracy	87.500%	<b>99.063%</b>
	Time	0.653	<b>0.658</b>
Mnist in Csv data set	Accuracy	100.0%	<b>100.00%</b>
	Time	2.094	<b>1.498</b>
Music Genre data set	Accuracy	98.611%	<b>99.566%</b>
	Time	0.758	<b>0.653</b>

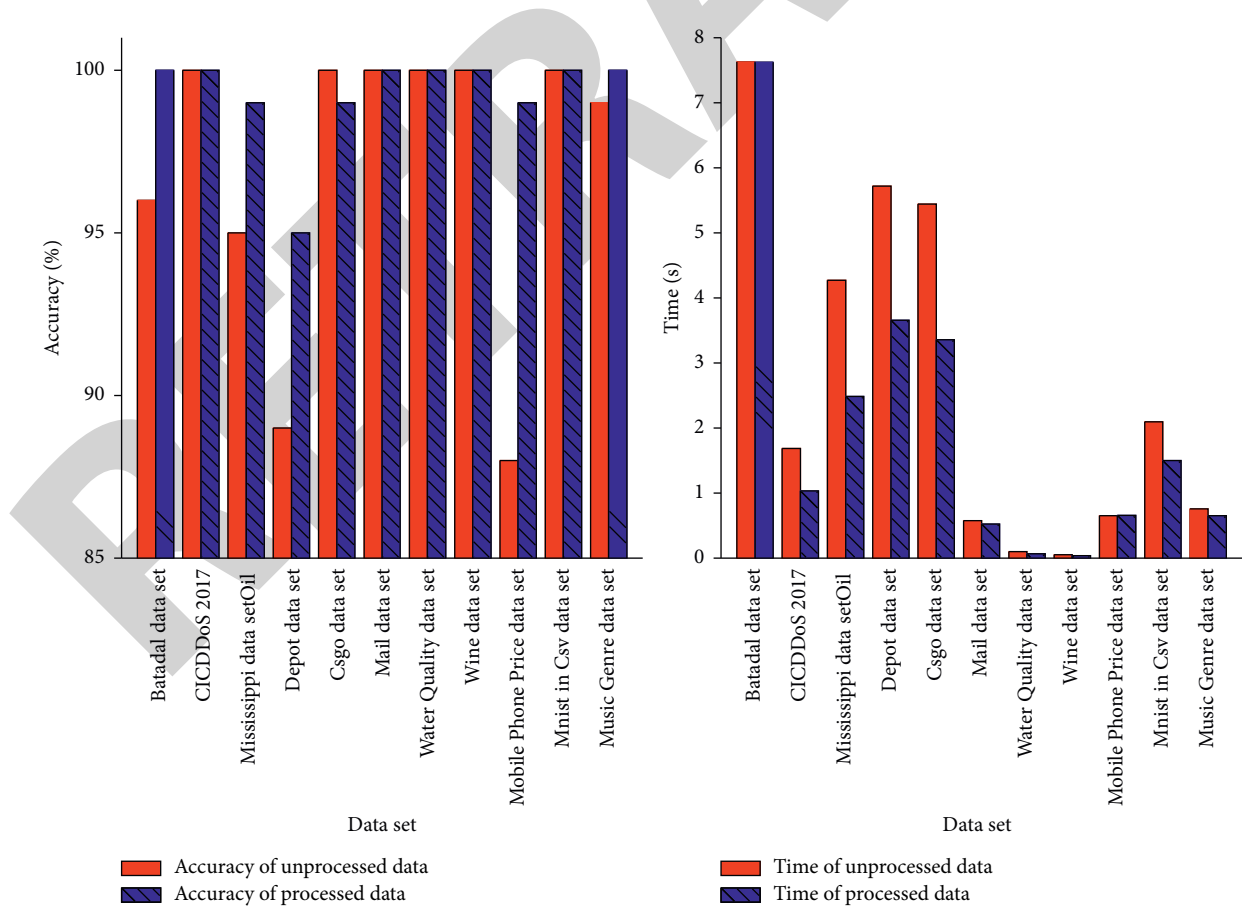


FIGURE 4: Comparison of data sets before and after feature sequencing algorithm.



higher than before or 100%. On the right is the comparison diagram of the average time of each generation. The sequencing algorithm of all data sets can reduce the calculation time.

## 5. Conclusion

This paper aims to design an adaptive algorithm to find the optimal feature sequencing algorithm for different data sets. Through a variety of evaluation indexes of the data set and decision tree algorithm, the appropriate feature sequencing algorithm is selected. The selected algorithm is used to sequence the features of the data set. Then, the sequenced data set is classified by neural network anomaly detection algorithm. This paper compares the effects of various feature sequencing algorithms on anomaly detection accuracy and training speed and verifies the effect of the algorithm selected by the adaptive method in this paper.

In this paper, 7 common feature sequencing algorithms are used and 11 public data sets in industrial control field and other fields are used for experiments. In the experimental results, the feature sequencing algorithms selected by this algorithm for all data sets are the algorithms with the highest accuracy and higher training speed than the average speed.

By comparing the original data set not processed by this algorithm with the processed 11 data sets in the anomaly detection algorithm, the results show that the accuracy of 30 generations of all data sets is improved by 2.568% on average and the average time of each generation of all data sets is shortened by 24.143%. This algorithm can effectively select the feature sequencing algorithm suitable for different data sets, improve the accuracy of anomaly detection, reduce the training time, and reduce the influence of feature distribution on the anomaly detection algorithm.

This paper mainly studies the adaptive algorithm applied to the data of industrial control systems. The selected experimental algorithm and index are highly targeted, so the algorithm has some limitations. The experimental results show that the accuracy of this algorithm is lower than that before processing in the experiment using Csgo data set. The next step is to improve the evaluation index system and increase the experimental data set to enhance the universality of the selection method.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported by BIPTACF-008.

## References

- [1] Y. Tao, L. Ning, and G. Zeng, "Review of industrial control systems security," *Computer Engineering and Applications*, vol. 52, no. 13, pp. 8–18, 2016.
- [2] P. A. S. Ralston, J. H. Graham, and J. L. Hieb, "Cyber security risk assessment for SCADA and DCS networks," *ISA Transactions*, vol. 46, no. 4, pp. 583–594, 2007.
- [3] S. Ding, Z. Wang, W. Kong, H. Yang, and G. Song, "Electrode regulating system modeling in electrical smelting furnace using recurrent neural network with attention mechanism," *Neurocomputing*, vol. 359, pp. 32–40, 2019.
- [4] F. Zhu, W. Wu, and Y. Fu, "A dual deep network based secure deep reinforcement learning method," *Chinese Journal of Computers*, vol. 42, no. 8, pp. 1812–1826, 2019.
- [5] J. Inoue, Y. Yamagata, Y. Chen, C. M. Poskitt, and J. Sun, "Anomaly detection for a water treatment system using unsupervised machine learning," in *Proceedings of the 2017 IEEE International Conference on Data Mining Workshops (ICDMW)*, pp. 1058–1065, New Orleans, LA, USA, November 2017.
- [6] V. Todd and M. Milos, "Computationally efficient neural network intrusion security awareness," in *Proceedings of the 2009 2nd International Symposium on Resilient Control Systems*, pp. 25–30, New York, August 2009.
- [7] L. Ondrej, V. Todd, and M. Milos, "Neural network based intrusion detection system for critical infrastructures," in *Proceedings of the 2009 International Joint Conference on Neural Networks*, pp. 1827–1834, Atlanta, GA, USA, June 2009.
- [8] W. Wang, J. Guo, Z. Wang et al., "Abnormal flow detection in industrial control network based on deep reinforcement learning[J]," *Applied Mathematics and Computation*, vol. 409, Article ID 126379, 2021.
- [9] A. Esra and B. Mustafa Gokce, "An ensemble-based semi-supervised feature ranking for multi-target regression problems," *Pattern Recognition Letters*, vol. 148, pp. 36–42, 2021.
- [10] P. Porizka, J. Klus, E. Kepes, J. Kaizer, and D. M. Hahn, "On the utilization of principal component analysis in laser-induced breakdown spectroscopy data analysis, a review," *Spectrochimica Acta Part B: Atomic Spectroscopy*, vol. 148, pp. 65–82, 2018.
- [11] T. Chen and H. Dong, "(n, d) Industrial anomaly intrusion detection using ant colony algorithm and deep reinforcement learning," *Journal of Chinese Computer Systems*, vol. 1-8, 2021, <http://kns.cnki.net/kcms/detail/21.1106.TP.20210319.1122.024.html>.
- [12] K. Ron and H. John George, "Wrappers for feature subset selection," *Artificial Intelligence*, vol. 97, no. 1-2, pp. 273–324, 1997.
- [13] S. Patrick, B. Christa, and A. S. Lothar, "Correlation coefficients: appropriate use and interpretation," *Anesthesia and analgesia: Journal of the International Anesthesia Research Society*, vol. 126, no. 5, pp. 1763–1768, 2018.
- [14] S. Yu, W. Tan, C. Zhang, Y. Fang, C. Tang, and D. Hu, "Research on hybrid feature selection method of power transformer based on fuzzy information entropy," *Advanced Engineering Informatics*, vol. 50, Article ID 101433, 2021.
- [15] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society: Series B*, vol. 58, no. 1, pp. 267–288, 1996.
- [16] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," *Journal of the Royal Statistical Society: Series B*, vol. 67, no. 2, pp. 301–320, 2005.

## Retraction

# Retracted: High-Performance Computational Recognition of Communication Signals Based on Bispectral Quadratic Feature Model

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] Y. Chen, R. Zhu, J. Guo, and F. Wang, "High-Performance Computational Recognition of Communication Signals Based on Bispectral Quadratic Feature Model," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 2773492, 12 pages, 2022.

## Research Article

# High-Performance Computational Recognition of Communication Signals Based on Bispectral Quadratic Feature Model

Yarong Chen , Rui Zhu, Jianxin Guo, and Feng Wang

School of Information Engineering, Xijing University, Xi'an, Shaanxi 710123, China

Correspondence should be addressed to Yarong Chen; 20170063@xijing.edu.cn

Received 25 January 2022; Accepted 25 March 2022; Published 26 April 2022

Academic Editor: Daqing Gong

Copyright © 2022 Yarong Chen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In response to the problems in the signal identification of radiation sources during the communication process, the bispectral quadratic feature model is applied to the identification algorithm for communication signals. According to the signal eigenvalues obtained from the bispectrum of the diagonal slices in the radiation source signals, the eigenvalues of the bispectrum diagonal slices can be extended from the frequency domain to the complex plane through the chirp-z operation in this paper, and the relevant data are obtained based on the bispectrum quadratic feature model of the signals by using the separation rules corresponding to the extended Babbitt distance. The bispectral quadratic feature model method is used to establish a sparse observation model, and the communication signal processing problem can be transformed into an estimation problem of signal motion parameters through the construction of a parametric database. At the same time, the high-resolution distance of communication signals is tested, and the communication signals are estimated by using the variational inference method. Finally, practical cases are analyzed, and the results indicate that the algorithm proposed in this paper can be used to identify different types of communication signals in accordance with simulated and measured data in the processing of communication signals in various environments, which has the certain anti-interference capacity to noise, can improve the identification rate of communication signals, and has verified the effectiveness and practicality of the algorithm proposed in this paper.

## 1. Introduction

The identification of communication signals is one of the key means of the communication industry at present, and it has been widely used in all walks of life and in the military field. As carriers of information data transmission, signals contain the features of the radiation source, which can be used to achieve accurate identification and detailed analysis of the communication signal, obtain the communication radiation source bispectrum features, and provide a basis for the subsequent communication signal radiation source effective identification [1, 2]. At present, for the purpose of implementing accurate identification of communication radiation source signals, generally selected communication signal bispectral features should comply with the features of the time shift without changes, the size without changes, and the

phase without changes. In the signal bispectrum analysis, based on compliance with the aforesaid conditions while also needing to maintain high noise immunity, it is widely applied in different industrial fields [3, 4]. As the bispectrum analysis process is relatively simple and the workload of operations is relatively small, it can maintain a high level of higher-order spectral analysis features. Hence, it is more extensively used in the field of communication signal processing to acquire the bispectral quadratic features of the signal complex diagonal tangent on the basis of in-depth rooted communication signal bispectral quadratic features. Due to the existence of different identification accuracy of different high-performance computing subpaths, the use of the bispectral quadratic feature model can be obtained in accordance with the feature vector of the communication signal in the results of high-performance computing. Finally,

the results of the experiment indicate that the proposed method is practical for the secondary identification algorithm of the communication signal features using the baroclinic distance criterion [5, 6]. As the communication signal has the innate advantages of all-day real-time, long-range role, high resolution, and so on, it can be used on a large scale in many fields such as high-resolution processing, feature extraction, signal classification, and so on, and the use of transmitting bispectral quadratic feature model can increase the acquired high-distance resolution. At the same time, it can also shorten the features of the receiver transient bandwidth, complete hardware reduction, and according to the subpulse carrier, frequency sequence can effectively reduce the observation time while effectively improving the anti-interference performance of the communication signals. Communication signals mainly indicate the signals of different factors (noise, no role of the signal) fused together, and the transmission path is more complex [7–9]. Even if a higher function of the communication signal is used, the signal results that can be obtained are not ideal. Hence, it is not easy to obtain a physical feature of the initial signal. However, deconvolution is a method to obtain the original signal in accordance with the feedback from the fused signal. This method is extensively used in different fields such as communication, communication signals, speech, and medicine. One of the more technologically advanced methods at present is the variational difference recovery method proposed by Chan et al. This calculation uses a partial differential gradient projection method for the Lagrange multiplier term to minimize the difference [10–12]. Its advantages are fast convergence, stability, and so on, especially for signals with steep edges. With regard to the regular dynamic adaptive calculation method, the numerical value of the difference is compared accordingly.

The bispectral quadratic feature model algorithm is applied to the issue of high-performance computing identification and analysis of communication signals in this paper by converting the estimation of motion parameters of communication signal identification into the evaluation problem of identification and analysis of reconstruction. This method focuses on the effective synthesis of low signal-to-noise ratio HR-RP based on the parameters of the communication signal identification motion signal in the search interval. In this way, it can quickly implement the estimation of the motion parameter impact under the premise of solving the HRRP reconstruction error under a low signal-to-noise ratio to identify the local optimal value according to the population update speed and finally accomplish the accurate processing of the communication signal motion parameters.

## 2. Models and Algorithm

**2.1. Bispectral Quadratic Feature Model.** In accordance with the setting of the carrier frequency sequence of the communication signal, it is possible to obtain the results of the bispectral quadratic feature model in accordance with the complete step FM signal within [13, 14]. In this way, the

whole stepper FM signal contains  $N$  pulses, while  $M$  pulses ( $M < N$ ) are extracted according to the sequence.

As the signal carrier frequency sequence of the bispectral quadratic feature model is set to  $f_{sm} = f_c + g(m)\Delta f$ , in which  $\Delta f$  stands for the pulse bandwidth of the signal and  $g$  stands for a subset in the interval  $[0:N-1]$ , then the bandwidth of the bispectral quadratic feature model can be expressed as  $B = N\Delta f$ . Through the detailed analysis, it can be known that the corresponding correlation accumulation time is no greater than the whole communication signal step FM signal. Then the carrier frequency sequence of the communication signal can be set in accordance with the environmental information so that it possesses the enhanced anti-interference capability of itself.

It is assumed that the communication signal transmits a total of  $k$ -th groups of sparse step frequency modulation (FM) signals, then the sparse step FM high-performance operation of the first group can be expressed as follows:

$$s_1(t) = \sum_{m=0}^{M-1} \text{rect}\left(\frac{t - mT_R - kMT_R}{T_p}\right) \cdot \exp(j\pi\gamma(t - mT_R - kMT_R)^2) \cdot \exp(j2\pi f_{sm}(t - mT_R - kMT_R)), \quad (1)$$

where  $t = \hat{t} + mT_R + kMT_R$  ( $m = 1, 2, \dots, M$ ) stands for the whole calculation time,  $\hat{t}$  stands for the fast time, and  $\text{tect}(u)$  stands for the corresponding rectangular window. If the value of  $\text{tect}(u)$  is 1 when  $|u| \leq 1/2$ , then the value of  $\text{tect}(u)$  will become 0, in which  $\gamma$  indicates the modulation frequency of the communication signal and  $T_p$  and  $T_R$  stand for the pulse width of the communication signal and the number of pulse repetition periods in turn, respectively.

If the communication signal contains multiple scattering points, then the coefficient of the backward scattering of the  $p$  ( $p = 1, 2, \dots, P$ ) scattering point can be expressed as  $\sigma_p$ , which constitutes the ‘‘Stop-Go’’ model. The corresponding time delay value  $\tau_p(t)$  of the scattering point  $p$  within the pulse of the communication signal can be left unchanged, and then there will be  $\tau_p(t) \approx \tau_p(t_{m,k})$ ,  $t_{m,k} = mT_R + kNT_R$ . At the same time, with regard to the scattering point of the communication signal, the  $m$ -th subpulse echo under the  $k$ -th group of sparse step FM high-performance operation can be expressed as follows:

$$s_2(\hat{t}, m, k) = \sigma_p \text{rect}\left(\frac{\hat{t} - \tau_p(t_{m,k})}{T_p}\right) \cdot \exp(j\pi\gamma(\hat{t} - \tau_p(t_{m,k}))^2) \cdot \exp(j2\pi f_{sm}(\hat{t} - \tau_p(t_{m,k}))) + \varepsilon(\hat{t}), \quad (2)$$

where  $\tau_p(t_{m,k}) = 2R_p(t_{m,k})/c$ ,  $R_p(t_{m,k})$  stands for the instantaneous slope distance between the  $p$ -th scattering point and the communication signal emission signal, in which  $c$  is the speed of light and  $\varepsilon(\hat{t})$  indicates the additive noise. If the accumulation angle  $\theta_{m,k}$  of the signal to be processed is



reduced,  $R_p(t_{m,k}) = R(t_{m,k}) + x_p \sin \theta_{m,k} + y_p$ , in which  $R(t_{m,k})$  stands for the instantaneous slope distance between the communication signal,  $(x_p, y_p)$  stands for the reference point by the instantaneous slope distance, and the scattering point  $p$  of the communication signal is denoted by the coordinate value in its plane. Thus, for a high-speed communication signal, the distance between the signal reference point and the communication signal at the time point  $t_{m,k}$  can be obtained as  $R(t_{m,k}) = \tau_R + v_R t_{m,k} + 1/2 a_R t_{m,k}^2$ , where  $r_R$  stands for the distance between the communication signal and the communication signal at the initial time,  $v_R$  stands for the radial velocity of the signal, and  $a_R$  stands for the radial acceleration of the communication signal. The time delay of the communication signal can be expressed as  $\tau_{ref}(t_{m,k}) = 2R_{ref}(t_{m,k})/c$ , where  $R_{ref}(t_{m,k}) = r_R + \hat{v}_R t_{m,k} + 1/2 \hat{a}_R t_{m,k}^2$ . In general, the estimation of  $v_R$  and  $a_R$  can be obtained in the tracking phase of the communication

signal. Then, for the echoes to be solved by line frequency modulation, it can be assumed that  $\hat{f} = \gamma(\hat{t} - 2R(t_{m,k})/c)$ ; thus, the communication signalization is converted to the following equation:

$$s_3(\hat{f}, m, k) = \sigma_p \text{rect}\left(\frac{\hat{f}}{\Delta f}\right) \exp\left(j \frac{4\pi}{c} (f_{sm} + \hat{f}) \Delta R\right) \cdot \exp\left(j \frac{4\pi}{c} (\Phi_P + \Phi_B)\right) + \varepsilon(\hat{f}), \quad (3)$$

where  $\Delta R = x_p \sin \theta_{m,k} + y_p$  and  $\Phi_P$  and  $\Phi_B$  stand for the phase error due to interpulse translation of the communication signal and the phase error due to interpulse translation of the pulse string, respectively. Thus,  $\Phi_P$  and  $\Phi_B$  can be expressed as follows:

$$\left. \begin{aligned} \Phi_P &= \Phi_1 + \Phi_2, \Phi_B = \Phi_3 + \Phi_4 + \Phi_5 \\ \Phi_1 &= m^2 \left( \frac{1}{2} \Delta a_R f_0 T_R^2 + \Delta f \Delta v_R T_R + k \Delta f \Delta a_R M T_R^2 + \frac{1}{2} \hat{f} \Delta a_R T_R^2 \right), \Phi_2 = m^3 \left( \frac{1}{2} \Delta a_R \Delta f T_R^2 \right) \\ \Phi_3 &= m \left( \begin{array}{c} f_0 \Delta v_R T_R + k \Delta f \Delta v_R M T_R + k \Delta a_R f_0 M T_R^2 + \hat{f} \Delta v_R T_R + \\ \frac{1}{2} \Delta a_R \Delta f (k M T_R)^2 + k \Delta a_R \hat{f} M T_R^2 \end{array} \right) \\ \Phi_4 &= k (f_0 \Delta v_R M T_R + \hat{f} \Delta v_R M T_R), \Phi_5 = k^2 \left( \frac{1}{2} \Delta a_R f_0 (M T_R)^2 + \frac{1}{2} \hat{f} \Delta a_R (M T_R)^2 \right) \end{aligned} \right\}, \quad (4)$$

where  $\Delta v_R$  stands for the residual velocity;  $\Delta v_R = v_R - \hat{v}_R$ , in which  $\Delta a_R$  stands for the residual acceleration; and  $\Delta a_R = a_R - \hat{a}_R$ .  $\Phi_1$  stands for the  $m$  quadratic phase term, which will generate the primary flap spreading in the phase of distance synthesis of the communication signal.  $\Phi_2$  stands for the tertiary phase term of  $m$ ; the communication signal distance image is synthesized as an asymmetric paraflap; and then this term can usually be expressed in the order of  $10^{-4}$ . Hence, it can be neglected.  $\Phi_3$  contains the coupling terms of both  $m$  and  $k$ , which will lead to a distance image shift and cause the bending of the envelope.  $\Phi_4$  stands for the primary phase term of  $k$ , which causes an azimuthal shift in the image of the communication signal and can also be neglected.  $\Phi_5$  stands for the quadratic phase term of  $k$ . The azimuthal pulse pressure will cause the main flap spreading of the communication signal. For the purpose of constructing the reconstruction algorithm with sparse HRRP efficiently, the echoes can be converted into discrete form. If the number of pulse sampling points of each communication signal is set to  $N_r$ , which complies with  $\bar{L} = M \cdot N_r$  at the same time, then the  $k$ -th group of echoes can be expressed as  $s_k = [s_{0,k}, \dots, s_{m,k}, \dots, s_{M-1,k}]_{1 \times \bar{L}}^T$ , in which  $s_{m,k} = [s_{0,m,k}, \dots,$

$s_{N_r-1,m,k}]_{1 \times N_r}$ ,  $s_{n,m,k} = \exp(j4\pi/c(f_{sm} + n_r/N_r \Delta f)(\Delta R + \Delta v_R t_{m,k} + 1/2 \Delta a_R t_{m,k}^2)) + \varepsilon(n_r)$ .

The remaining transmission process parameters of the communication signal are introduced into the database, and then the sparse observation model of the communication signal can be expressed as follows:

$$s_k = D_k(\Delta v_R, \Delta a_R) \theta_k + n, \quad (5)$$

where  $D_k(\Delta v_R, \Delta a_R) \in \mathbb{C}^{\bar{L} \times \bar{L}}$  stands for the database matrix corresponding to the  $n$ -th echo,  $L = N \cdot N_r$ ,  $\theta_k \in \mathbb{C}^{L \times 1}$  stands for the HRRP corresponding to the  $k$ -th echo, and  $n$  stands for the noise vector. Thus,  $d_l^k$  can be obtained as follows:

$$d_l^k(\Delta \hat{v}, \Delta \hat{a}) = f_l \odot g_k \odot h_k, \quad (6)$$

where  $\odot$  stands for the inner product,  $f_l$  stands for the  $l$ -th column of the database  $F$ ,  $F = [F_0, \dots, F_{M-1}]_{L \times \bar{L}}^T$ ,  $g_k = [g_{0,k}, \dots, g_{M-1,k}]_{\bar{L} \times 1}^T$ , and  $h_k = [h_{0,k}, \dots, h_{M-1,k}]_{L \times 1}^T$ . In accordance with the  $m$ -th subpulse of the signal, it can be known that

$$F_m = \begin{bmatrix} W_m(0, 1) & \cdots & W_m(0, L-1) \\ \vdots & \vdots & \vdots \\ W_m(N_r-1, 0) & \cdots & W_m(N_r-1, L-1) \end{bmatrix}_{N_r \times L}, \quad (7)$$

$$g_{m,k} = [g_{m,k}(0) \cdots g_{m,k}(N_r-1)]^T, g_{m,k}(n_r) = \exp\left(j \frac{4\pi}{c} \left(f_{sm} + \frac{n_r}{N_r} \Delta f\right) \Delta v_R t_{m,k}\right),$$

$$h_{m,k} = [h_{m,k}(0) \cdots h_{m,k}(N_r-1)]^T,$$

$$h_{m,k}(n_r) = \exp\left(j \frac{2\pi}{c} \left(f_{sm} + \frac{n_r}{N_r} \Delta f\right) \Delta a_R t_{m,k}^2\right). \quad (8)$$

As the signal reconstruction method based on the bispectral quadratic feature model can operate on the real matrix, it can be expressed as follows according to equation (8):

$$y = \Phi\theta + \varepsilon. \quad (9)$$

In accordance with expression (9), the probability map model is established by using the Gamma-Gaussian algorithm [15, 16]. If the acquired noise  $\varepsilon$  is Gaussian white noise with 0 mean, then the probability of the signal echo  $y$  thus obtained is also Gaussian distributed, and the probability of  $\varepsilon$  distributed with  $y$  can be obtained as follows:

$$p(\varepsilon) = N(\varepsilon|0, \alpha^{-1}I), \quad (10)$$

$$p(y|\theta, \alpha) = N(y|\Phi\theta, \alpha^{-1}I),$$

where  $\alpha$  stands for the noise accuracy. The Gamma-Gaussian prior is introduced to the sparse vector  $\theta$ , and the following can be obtained:

$$p(\theta|\Lambda) = N(\theta|0, \Lambda^{-1}), \quad (11)$$

where the accuracy matrix  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_d)$  stands for the dimensional diagonal array, in which  $D = 2L$ . The accuracy parameters  $\lambda_d$  ( $d = 1, \dots, D$ ) and  $\alpha$  comply with the Gamma distribution, and the following can be obtained:

$$p(\lambda_d) = \text{Gamma}(\lambda|v_1, v_2), \quad (12)$$

$$p(\alpha) = \text{Gamma}(\alpha|v_3, v_4).$$

The product of the distributions according to the probability model can be obtained as follows:

$$p(y, \theta, \Lambda, \alpha) = p(y|\theta, \alpha)p(\theta|\Lambda)p(\alpha) \prod_d p(\lambda_d). \quad (13)$$

On the other hand, the posterior distribution of the random variables can be expressed as the joint distribution of the variables divided by the marginal distribution  $p(y)$ , and thus, the following can be obtained:

$$p(\theta, \Lambda, \alpha|y) = \frac{p(y, \theta, \Lambda, \alpha)}{p(y)}. \quad (14)$$

It is highly difficult to calculate the posterior distribution based on the model directly. Thus, high-performance

computing is carried out to resolve the approximate posterior distribution so as to enable the HRRP synthesis under the low signal-to-noise ratio (SNR) conditions.

## 2.2. Identification of High-Performance Computing of Communication Signals

### 2.2.1. Processing of Signals with Large Angle and Low Signal-to-Noise Ratio.

In the research on the performance of communication signals, the parameters related to the communication signals are analyzed, with the main parameters as follows:

- (1) Insertion loss (IL): The loss of signal power due to the addition of a device in a transmission line or fiber is denoted by decibels (dB). When the power transmitted to the load before insertion is  $P_{in}$  and at the same time the power received by the load after insertion is  $P_{\mathcal{L}}$ , then the equation for insertion loss in dB can be obtained as follows:

$$IL = 10 \log \frac{P_{in}}{P_{\mathcal{L}}} = -10 \log (1 - |\Gamma_{in}|), \quad (15)$$

where  $|\Gamma_{in}|$  stands for the reflection coefficient, which can be expressed as follows in accordance with the connection between IL and  $\mathcal{S}_{21}$ :

$$IL = 10 \log \frac{1}{|\mathcal{S}_{21}|^2} = -10 \log |\mathcal{S}_{21}|^2. \quad (16)$$

The relationship between  $P_{\mathcal{L}}$ ,  $P_{in}$ , and  $\mathcal{S}_{21}$  in equation (16) is as follows:

$$\frac{P_{in}}{P_{\mathcal{L}}} = \frac{1}{|\mathcal{S}_{21}|^2}. \quad (17)$$

- (2) Bandwidth represents the selected range of the signal, that is, the width of the spectrum to be passed, the unit of the frequency range width, which is expressed in Hz, with the following equation:

$$BW^{X\text{dB}} = f_H^{X\text{dB}} - f_L^{X\text{dB}}. \quad (18)$$

When the value of  $X$  is 3, 1, or 0.5,  $f_H$  and  $f_L$  stand for the frequencies on both sides of the passband

when the reduction in the values of insertion loss on the left and right sides of the center frequency is  $X$  (dB).

- (3) In-band fluctuation refers to the size of the fluctuation in the response amplitude of the passband, that is, the difference between the maximum and minimum values of the response amplitude. In the design process of communication signals, the smaller the in-band fluctuation, the better the performance of the communication signals.
- (4) Return loss (RL) refers to the functional loss in the signal returned or reflected by the discontinuity in the transmission line. Discontinuity may not comply with the terminating load or the equipment inserted into the line, which is expressed in decibels (dB). The reflection system of the communication signal is also referred to as the reflection loss. The equation for return loss is as follows:

$$RL = -10 \log \left[ \frac{VSWR - 1}{VSWR + 1} \right]^2. \quad (19)$$

When VSWR represents the voltage standing wave ratio, the magnitude of the return loss is related to the standing wave ratio (VSWR) and the reflection coefficient ( $\Gamma$ ). The increased return loss corresponds to a lower VSWR. Return loss is an index that measures how well the equipment matches the line. If the return loss is high, it means that the match is good. Generally, in the design of a communication signal, a high return loss and a low insertion loss mean that the performance of the communication signal is relatively good.

- (5) VSWR refers to the value of the impedance matching between the load and the transmission line or waveguide. A larger value of the VSWR indicates a higher degree of matching. The VSWR stands for the ratio of the amplitude of the partial standing wave at the wave web (maximum value) to the amplitude of the node along the line (minimum value).
- (6) The quality factor  $Q$  stands for the ability of the communication signal to allow separation of the neighboring frequencies in the signal, and its equation is as follows:

$$Q = \frac{f_o}{BW}. \quad (20)$$

When  $f_o$  represents the center frequency of the communication signal, BW stands for the 3 dB bandwidth. A larger value of  $Q$  indicates that the device has higher working stability, stronger frequency selectivity, lower loss, but narrower band; on the contrary, a smaller value of  $Q$  value indicates that the device's operating stability is lower, the frequency selectivity is weaker, the loss is relatively large, but the frequency band is wider. When the value of  $Q$  is increased in the design, the higher the  $Q$  value is, the stronger the performance of the device is.

- (7) Resonant frequency: With regard to the communication signals, there are multiple ways to resolve the resonant frequency. Thus, there are also various factors affecting the resonant frequency. Common factors affecting the resonant frequency include the structure of the communication signals, the shape and the resonant mode, and so on. In practice, four methods are often used to resolve the resonant frequency: the electro-nano method, the total set parameter method, the field solution method, and the phase method.

The indicators mentioned above are used to measure the performance of a communication signal. In practice, it is not required to set the limit for all indicators, and some of the indicators can be optimized according to the actual demand.

*2.2.2. Processing Result of Measured Signals.* The transfer function of a communication signal is a mathematical formula that expresses the frequency response features of the communication signal. With regard to the classical two-port communication signal, the transmission function equation is as follows:

$$|S_{21}(j\omega)^2| = \frac{1}{1 + \varepsilon^2 \mathcal{F}_n^2(\Omega)}, \quad (21)$$

where  $\varepsilon$  stands for the ripple coefficient,  $\mathcal{F}_n(\Omega)$  stands for the functional characteristic of the low-pass prototype, and  $\Omega$  stands for the frequency variable.

The features of communication signals are represented through the frequency response characteristics. With regard to the classification of communication signals as mentioned at the beginning of this paper, they can be divided into four types (i.e., low pass, high pass, band pass, and band resistance) according to the frequency response characteristics. The high pass, band pass, and band resistance can be obtained through the low-pass prototype of the frequency and components.

In the communication process, radiation source noise is mainly generated by the transmitter noise. In general, it refers to the amplitude of the communication signal, frequency, and pulse width and repetition frequency together causing abnormal changes, that is, the stability of the communication process signal resulting in radiation source noise. Signal instability can be roughly divided into two categories: regular instability and random instability. Regular instability is mainly due to inadequate power filtering, mechanical jitter, and so on; random instability is the noise generated by the transmitter tube and modulation pulse random jitter generated.

At present, the gradual improvement of communication signal protocol will help to improve the stability of transmitter. Thus, with the large-scale use of the main vibration amplification type of transmitter, the noise generated can mainly be divided into the following three aspects:

- (1) According to the amplitude frequency characteristics and phase frequency characteristics of the system, the cause of frequency-domain distortion can be analyzed.



- (2) The top of the modulated pulse generates vibration; the top begins to fall with the power supply fluctuations of the transmitter due to the signal corresponding to the parasitic phase or the amplitude caused by the time-domain distortion phenomenon.
- (3) The main control oscillator has insufficient frequency stability and phase stability. The different circuits or devices used for different communication transmitters result in different transmitter noises for various communications.

Hence, the uncoordinated modulation within the communication signal pulse is due to the transmitter noise. As the noise is generated due to various kinds of parasitic modulation. The communication signals present differences in signal features.

The noise output from the communication transmitter indicates more non-Gaussian and non-linear features, and in the bispectral analysis, the signal amplitude and phase information are maintained. At the same time, the effect of Gaussian non-color noise on non-Gaussian signals bispectrum is completely suppressed, which can be used for the extraction of unconscious modulation features. Thus, the following concept of bispectrum is obtained accordingly.

It is assumed that the high order cumulant  $c_{kx}(\tau_1, \tau_2, \dots, \tau_{k-1})$  is absolutely summable, that is, the following can be obtained:

$$\sum_{\tau_1=-\infty}^{\infty} \cdots \sum_{\tau_{k-1}=-\infty}^{\infty} |c_{kx}(\tau_1, \dots, \tau_{k-1})| < \infty. \quad (22)$$

The  $k$ -th-order spectrum is defined as the  $(k-1)$ -order discrete Fourier transform of the  $k$ -th-order cumulants, that is, the following can be obtained:

$$S_{kx}(\omega_1, \omega_2, \dots, \omega_{k-1}) = \sum_{\tau_1=-\infty}^{\infty} \cdots \sum_{\tau_{k-1}=-\infty}^{\infty} c_{kx}(\tau_1, \dots, \tau_{k-1}) \times \exp[-j(\omega_1 \tau_1 + \cdots + \omega_{k-1} \tau_{k-1})]. \quad (23)$$

Thus, the bispectrum, that is, the third-order spectrum, can be defined as follows:

$$B_x(\omega_1, \omega_2) = \sum_{\tau_1=-\infty}^{\infty} \sum_{\tau_2=-\infty}^{\infty} c_{3x}(\tau_1, \tau_2) \exp[-j(\omega_1 \tau_1 + \omega_2 \tau_2)]. \quad (24)$$

The bispectrum of the signal including phase noise can be obtained from the bispectrum estimation. The discrete noise signal obtained by Scout is represented by  $X(n) = s(n) + W(n)$ , where  $w(n)$  represents the Gaussian white noise signal,  $s(n)$  includes the signal of non-Gaussian noise output from the transmitter, and  $w(n)$  and  $s(n)$  are independent of each other. If the cumulative quantity is solved three times for  $x(n)$ , the following can be obtained:

$$c_{3x}(\tau_1, \tau_2) = E\{[s(n) + w(n)][s(n + \tau_1) + w(n + \tau_1)] \cdot [s(n + \tau_2) + w(n + \tau_2)]\}. \quad (25)$$

Equation (4) is expanded and combined to obtain the following:

$$c_{3x}(\tau_1, \tau_2) = c_{3s}(\tau_1, \tau_2) + c_{3w}(\tau_1, \tau_2) + E[w(n)][c_{2x}(\tau_1) + c_{2x}(\tau_2) + c_{2s}(\tau_2 - \tau_1)] + E[s(n)][c_{2w}(\tau_1) + c_{2w}(\tau_2) + c_{2w}(\tau_2 - \tau_1)]. \quad (26)$$

As long as the mean value of the signal and noise is zero, the following can be obtained:

$$c_{3x}(\tau_1, \tau_2) = c_{3s}(\tau_1, \tau_2) + c_{3w}(\tau_1, \tau_2). \quad (27)$$

As  $w(n)$  is a Gaussian noise signal,  $c_{3w}(\tau_1, \tau_2)$  can be excluded from the calculation. Thus, it can be known that the communication letter can be used to eliminate the white noise after the third-order accumulation, then the bispectrum can be determined by  $c_{3s}(\tau_1, \tau_2)$ , that is, the following can be obtained:

$$c_{3x}(\tau_1, \tau_2) = c_{3s}(\tau_1, \tau_2) = E\{s(n)s(n + \tau_1)s(n + \tau_2)\}. \quad (28)$$

In accordance with the above analysis, it can be observed that the evaluated bispectral features are mainly composed of the features of the signal itself and non-Gaussian noise. Hence, the bispectrum in the communication signal is assessed mainly based on the features specific to the signal itself, and it is also possible to obtain the features specific to various communications.

However, if the two-dimensional function can adopt the full-duplex spectrum as the signal feature to produce a two-dimensional template for matching, the number of operations can be excessively high, which will not comply with the high standard requirements for signal radiation source identification. The key to solving this issue lies in the introduction of high-performance computing bispectrum; in equation (28), the two-dimensional bispectrum is converted into a one-dimensional function. However, the high-performance computing bispectrum has the following defects:

- (1) The implementation of high-performance computing bispectrum is often a high-performance computation along each path. However, the secondary features obtained by this mode of computation are not consistent for the results to be recognized, and some of the bispectral points have relatively less effect on the results of the recognized targets and are subordinate to the ordinary bispectrum.
- (2) If there is a cross-term in the initial observed signal, the high-order accumulation calculated by using the multi-correlation function will lead to the result that the cross-term becomes more complicated. As the cross-term is generated by a random distribution, it is impossible to eliminate the cross-term based on the determined calculation method.

For the purpose of extracting the secondary features of the bispectrum as the features of the bispectrum and eliminating or decreasing the defects that occur in the process of high-performance computation of the

bispectrum, the characteristic parameters that are available for communication signals can be obtained with optimal separability in the secondary features of the bispectrum. As a result, it can effectively solve many problems such as cross-terms caused by ordinary bispectral points and high-performance computing.

In order to make the communication signal meet the actual performance characteristics, it is necessary to design an appropriate chamfer size in the design process. When the ring size becomes larger, the upper and lower passband edges move along the transmission zero point, and the attenuation pole becomes smaller. The intermediate frequency of the communication signal tends to be a higher frequency, which changes the frequency characteristics of the communication signal. With the shortening of the distance between patch resonances and the smaller the passband bandwidth, the attenuation point will first become smaller and then become larger. Therefore, HFSS software is used to optimize the structural parameters of the communication signal, and finally, the structural dimensions of the communication signal are obtained, as shown in Table 1.

By selecting a vector analysis instrument to test the communication signal, Figure 1 is the simulation test diagram corresponding to the communication signal. From Figure 1, it can be seen that the intermediate frequency of double passband communication signal is 3.8 GHz and 5.9 GHz, and the bandwidth corresponding to 3 dB is 13.5% and 6.2%, respectively. The minimum insertion loss in broadband is 0.7 dB and 1.3 dB in turn, and the return loss in the passband is not higher than 60 dB. In addition, there are zero points at the edges of the upper and lower passbands, which greatly improves the suppression characteristics of the stopband. The experimental data are in good agreement with the simulation data.

Since the conjugate prior is adopted in this paper for probabilistic modeling, the closed-form solution of the posterior distribution can be obtained directly by carrying out high-performance computing [17]. It is assumed that  $X$  stands for the observed data and  $w$  stands for the set of random variables; then the expression of the approximate posterior distribution can be obtained as follows:

$$\ln q(w_j) = E_{q_i \neq j} [\ln q(X, w)] + \text{const}, \quad (29)$$

where const. stands for a constant that guarantees the normalization of the posterior distribution. In the subsequent section, in accordance with the equation (18) and the probability model established, the approximate posterior distributions of the variables  $\theta$ ,  $\alpha$ , and  $\Lambda$  are solved for the sparse reconstruction by using the high-performance computing. In accordance with the mean field assumption, equation (14) is given in the form of the posterior multiplication of  $\theta$ ,  $\alpha$ , and  $\Lambda$ .

$$p(\theta, \Lambda, \alpha | y) \approx q(\theta)q(\Lambda)q(\alpha) = q(\theta) \prod_d q(\lambda_d)q(\alpha). \quad (30)$$

The steps for the  $\theta$  solution are described as follows:

TABLE 1: Structural parameters of level 1 UNICOM signal.

Structural parameters	Size (mm)
$a$	24
$b$	3
$c$	9
$d$	5
$e$	3
$f$	1.77
$g$	1.5
$L$	2

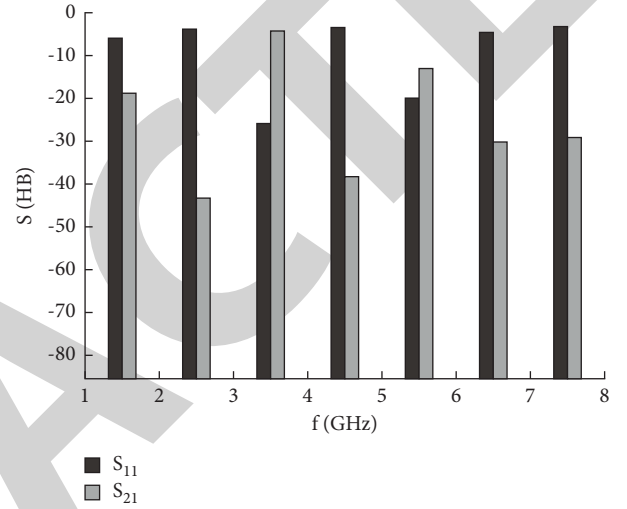


FIGURE 1: Simulation change diagram of communication signal.

*Step 1.* Initialization. Given  $\Phi$ ,  $v_1$ ,  $v_2$ ,  $v_3$ ,  $v_4$ ,  $\Lambda$ , and  $\alpha$  are initialized, the termination threshold is  $\eta_1$ , and the number of terminations is  $G_2$ .

*Step 2.* Update the variable  $\alpha$ , and its approximate posterior distribution is the Gamma distribution according to equation (14), as shown in the following equation:

$$q(\alpha) = \text{Gamma}(\alpha | a', b'), \quad (31)$$

where  $a' = v_3 + \bar{L}/2$  and  $b' = v_4 + 1/2 E_{q(\theta)} [(y - \Phi\theta)^T (y - \Phi\theta)]$ , in which  $\bar{L}$  stands for the length of  $y$ . At this point, the expectation of  $\alpha$  can be obtained as follows:

$$E_{q(\alpha)} [\alpha] = \frac{a'}{b'}. \quad (32)$$

*Step 3.* Update the variable  $\Lambda$ , and the approximate posterior distribution of the  $d$ -th element  $\lambda_d$  is the Gamma distribution, as shown in the following equation:

$$q(\lambda_d) = \text{Gamma}(\lambda_d | e'_d, f'_d), \quad (33)$$

where  $e'_d = v_1 + 1/2$  and  $f'_d = v_2 + 1/2 E_{q(\theta)} [\theta_d^2]$ . Thus, the expectation of  $\lambda_d$  can be obtained as follows:

$$E_{q(\lambda_d)} [\lambda_d] = \frac{e'_d}{f'_d}. \quad (34)$$

*Step 4.* Update the variable  $\theta$ , and its approximate posterior distribution is the Gaussian distribution, as shown in the following equation:

$$q(\theta) = N(\theta|\mu', \Sigma'), \quad (35)$$

where  $\Sigma' = (E_{q(\Lambda)}[\Lambda] + E_{q(\alpha)}[\alpha]\Phi^T\Phi)^{-1}$  and  $\mu' = E_{q(\alpha)}[\alpha]\Sigma'\Phi^T y$ . At this time, the following can be obtained:

$$E_{q(\theta)}[(y - \Phi\theta)^T(y - \Phi\theta)] = (y - \Phi\mu')^T(y - \Phi\mu') + \text{trace}(\Phi^T\Phi\Sigma'). \quad (36)$$

*Step 5.* Repeat Steps 2 to 4, and the iteration is suspended when the relative change of  $\hat{\theta}$  from the previous estimate does not exceed the threshold value  $\eta_1$ .

In the following section, the accuracy requirement for  $(\Delta\hat{v}_R, \Delta\hat{a}_R)$  is analyzed according to equation (4). With regard to the HRRP synthesis, the quadratic phase term  $\Phi_1$  can lead to distance image spreading. Hence, it is required that the change in the coherent accumulation time  $\Phi_1$  of the burst should be less than  $\pi/2$ . Thus, the following can be obtained:

$$\begin{aligned} |\Delta\hat{v}| &\leq \frac{c}{8M^2\Delta f T_R}, \\ |\Delta\hat{a}| &\leq \frac{c}{4f_0M^2T_R^2}. \end{aligned} \quad (37)$$

In practice, equation (26) can usually be met. Hence, the effect of  $\Delta\hat{a}$  on HRRP synthesis is neglected. For the purpose of eliminating the envelope bending caused by the first-order phase term field of  $\Phi_3$ , it is required that the envelope shift induced by the residual velocity and the residual acceleration during the processing observation time should not exceed the specified range. Thus, the following can be obtained:

$$\begin{aligned} |\Delta\hat{v}| &\leq \frac{c}{4M\Delta f (KMT_R)}, \\ |\Delta\hat{a}| &\leq \frac{c}{2M\Delta f (KMT_R)^2}. \end{aligned} \quad (38)$$

It should be noted that the linear phase of  $\Delta\hat{v}$  generated between the bursts leads to the shift and bending of the distance image and  $\Delta\hat{a}$  results in the widening and scattering of the azimuthal main flap.

For the purpose of obtaining excellent azimuthal focus, it is required that the peak reduction of the azimuthal image caused by the residual acceleration during the processing observation time should be no more than 3 dB; then the following can be obtained:

$$|\Delta\hat{a}| \leq \frac{7c}{4f_0(KMT_R)^2}. \quad (39)$$

Hence, azimuthal focusing requires higher accuracy of the acceleration estimation. Finally, for the purpose of obtaining a well-focused image, the signal residual velocity

and the residual acceleration should comply with equations (36) and (37). The subsequent experiments indicate that the algorithm proposed in this paper can comply with the requirement for estimation accuracy.

### 3. Simulation Experiment and Analysis

In this paper, the effectiveness of the proposed parameter estimation algorithm in the transmission process and the high resolution processing algorithm is verified based on the simulation data. The echoes of the satellite scattering point model (as shown in Figure 2 below) are generated in accordance with the parameters set out in Tables 2 and 3. It is assumed that the residual velocity and residual acceleration of the signals are 9 m/s and 1 m/s<sup>2</sup>, respectively. The bispectral quadratic feature model contains a total of 128 bursts, and each burst contains 64 randomly selected pulses from 80 consecutive full-band pulses (waveform 1). Through the addition of the complex Gaussian white noise to the echo, the echo signal-to-noise ratio can be increased from 0 dB to 15 dB in 5 dB steps. For each signal-to-noise ratio, 25 independent trials with different noise states are carried out with the number of genetic algorithm populations set to 40 and the number of genetic terminations set to 20. The algorithm proposed in this paper is compared with the PSO algorithm based on the parametric database (algorithm 1). In the comparison, the distance image entropy is weighted with the average distance image entropy as a signal function in accordance with algorithm 1.

The variation of the signal residual velocity and the acceleration estimation errors with the change in the SNR are shown in Figure 3. It can be observed that the estimation error of algorithm 1 at a low SNR is relatively large and fails to comply with equations (26) and (27). This is caused by the relatively large reconstruction error based on the OMP algorithm at the low SNR conditions. With the increase in the signal-to-noise ratio, the estimation error of algorithm 1 is decreased. However, the estimation results still fail to comply with the accuracy requirements and can result in image scattering. Compared with algorithm 1, the algorithm proposed in this paper is robust at any signal-to-noise ratio conditions. In addition, the errors of the residual velocity and residual acceleration estimation are  $1 \times 10^{-2}$  m/s and  $5 \times 10^{-3}$  m/s<sup>2</sup>, respectively, which can meet the estimation accuracy requirements of equations (26) and (27). The superior performance of the proposed algorithm in the motion parameter estimation under the low signal-to-noise ratio conditions is attributed to the following aspect: (1) the Gamma-Gaussian prior-based reconstruction algorithm that can implement the reconstruction of HRRP with high accuracy and (2) the excellent global optimization capability of the genetic algorithm.

As the estimation error of algorithm 1 at the signal-to-noise ratio of 0 dB is excessively large that the focusing processing cannot be implemented, for the purpose of a fair comparison, the database is established here in accordance with the motion parameter estimates obtained based on the proposed algorithm; then the OMP, GD, and high-performance calculations are carried out to resolve equation

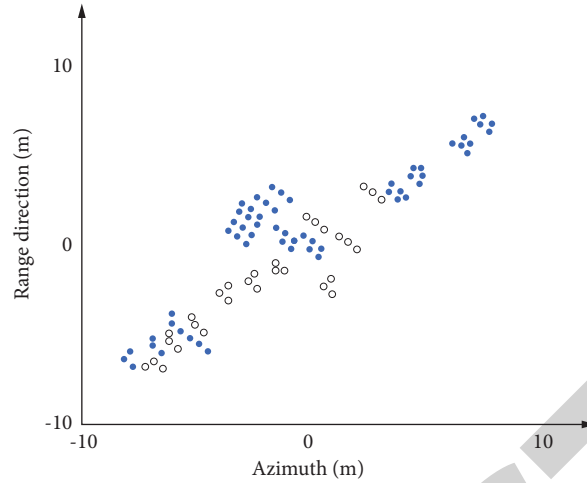


FIGURE 2: Distribution of the signal scattering points.

TABLE 2: Comparison of the operation complexity.

Reconstruction algorithm	OMP	GD	BQCM
Operational complexity	$\mathcal{O}(k_0 L^2)$	$\mathcal{O}(L^3)$	$\mathcal{O}(L^3)$

TABLE 3: Parameters of the communication signal system.

(GHz) $f_c$	PRF (KHz)	$T_R$	B (MHz)	$\Delta f$
10	6.4	20 ms	800	10 MHz

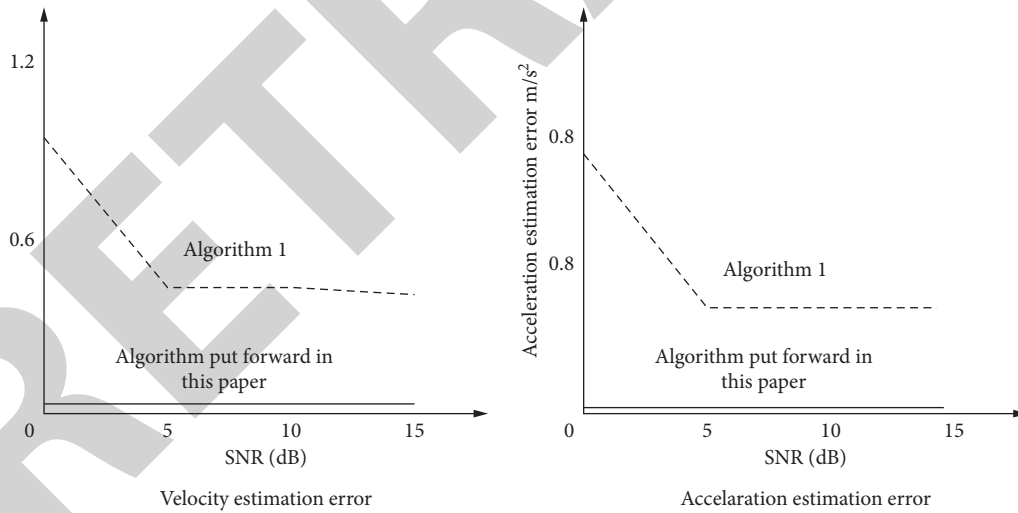


FIGURE 3: Variation curve of the velocity acceleration estimation error with the change of SNR: (a) velocity estimation error and (b) acceleration estimation error.

(9); and the images thus obtained are shown in Figure 4. From Figure 4, it can be observed that the algorithm proposed in this paper corresponds to the processing results with clear contours, fewer false points, and a better focusing effect than the other methods. In particular, the proposed algorithm can be used to describe the details of the solar sail panel more properly. In addition, the image entropy corresponding to the three methods can be obtained as 0.3310, 0.2935, and 0.2915.

For the purpose of fully verifying the proposed algorithm put forward in this paper, the stability of the three algorithms is further compared. The signal identification effect in the presence of external interference is analyzed accordingly. Figure 5 below shows the corresponding time-domain waveforms of the signals.

In the process of high-performance computing identification of communication signals, it is impossible to ensure that the reception length and quality of the target signals and

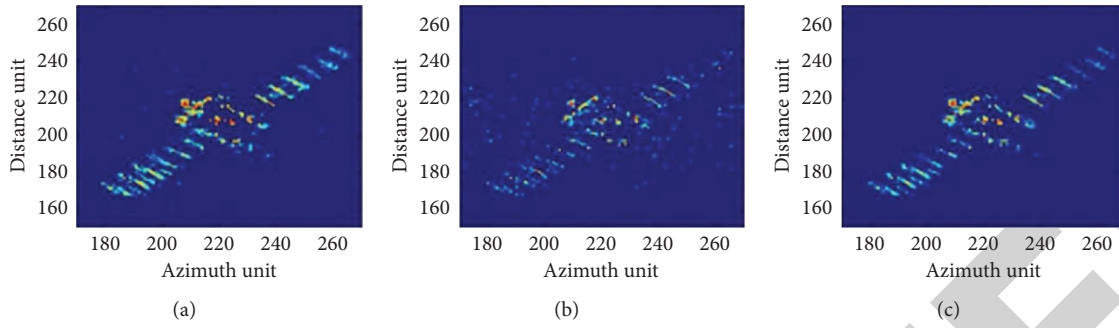


FIGURE 4: Processing results of the simulation data waveform 1: (a) OMP, (b) GD, and (c) VBI.

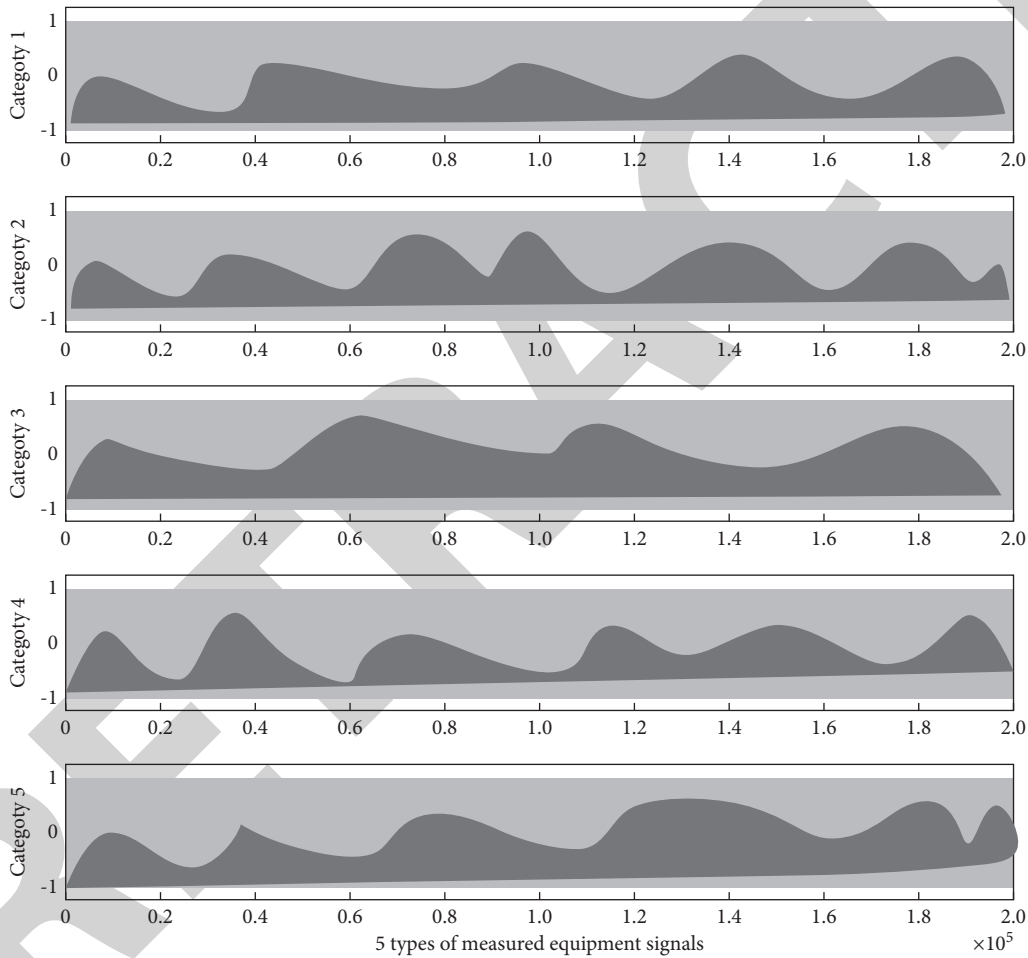


FIGURE 5: Time-domain waveform of the actual measured signal.

the computing under a certain number of conditions can present the stability of features, which is taken as an essential evaluation criterion for the practicality of the algorithm. Excellent methods for feature value extraction should not be constrained by the number of training, and such methods should have good robust performance when they are applied. After taking this point into full consideration, the identification accuracy of the training samples by using the target experimental signals is studied accordingly. The number of signal samples in each segment contains 500 points. For each class, the number of signal test samples is set

to 100, and the number of training samples is adjusted to 50, 100, 150, and 200, respectively. The experiment is carried out 30 times. The mean value and variance of the identification rates obtained based on the three methods are recorded in turn, which are shown in Table 4.

In the empirical testing, different modulation methods are used to increase the number of training samples as shown in Table 1, which has effectively improved the identification rate of the proposed algorithm. As a result, the identification performance starts to be stabilized. When the real number identification is compared by using the three



TABLE 4: Effect of different sample sizes on the identification results based on the three methods.

Experimental signal	$N_r$ number	BSB		ISIB		Method put forward in this paper	
		Mean value (%)	Variance ( $10^{-4}$ )	Mean value (%)	Variance ( $10^{-4}$ )	Mean value (%)	Variance ( $10^{-4}$ )
1, 2, 3, 4, 5	50	73.05	14.41	81.04	4.48	85.02	3.62
	100	74.65	8.36	81.21	3.98	86.06	2.38
	150	76.68	4.21	83.75	2.62	86.75	1.32
	200	77.15	6.53	84.27	3.14	87.21	1.54

methods, the method proposed in this paper has the highest identification rate for the real signals measured, followed by the ISIB method, and the BSB method is the least effective. When  $N_r$  is increased from 50 to 200, the identification rate of the distance selected bispectrum of the BAS is improved by about 4%, while the ISIB and the method proposed in this paper show the improvement of about 3% and 2%, respectively. The results of the experiment indicate that the algorithm put forward in this paper has excellent stability, the identification results are less affected by the number of training samples, and the model thus established has good robustness, which suggests that the proposed algorithm has relatively strong practicality.

#### 4. Conclusion

On the basis of the research on the fine features in the bispectral analysis of individual identification of radiation sources, the bispectral secondary features extracted are optimized by using the extended Babbitt distance criterion in accordance with the in-depth analysis of the bispectral features of communication signals diagonally sliced. This method can be used to implement the bounded deviation and identify the observation error values quickly and perform the complexity calculation of variational parameters effectively. The experiments indicate that the method is effective. The information on the original signals can also be recovered in a low signal noise background. The statistical features of communication scattering points and noise are used. Through practical analysis, the experimental results indicate that the method proposed in this paper has reduced the usage time. In addition, it can be applied to multiple types of communication signals. The bispectral features thus obtained always have excellent robustness at a low signal-to-noise ratio (SNR). When the signal-to-noise ratio is 0, the identification rate of high-performance computing of communication signals can achieve more than 90%. However, in general, a series of individual identification methods based on bispectral analysis have the common issue of a relatively low identification rate. Taking into account the subtle features in the other aspects of the signals, a feature vector is formed, which can further improve the practical effectiveness of the algorithm proposed in this paper.

#### Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

#### Conflicts of Interest

The authors declare that there are no conflicts of interest.

#### Acknowledgments

This research study was sponsored by (1) the Research Foundation of Xijing University, and the project number is XJ200105; (2) the Research Foundation for Talented Scholars of Xijing University, and the project numbers are XJ19B01, XJ20B01, and XJ17B06; (3) the foundation of Shaanxi Key Laboratory of Integrated and Intelligent Navigation, and the project number is SKLIIN-20190102; (4) Natural Science Foundation of Shaanxi Province, and the project numbers are 2021JM-537 and 2021GY-341. The authors would like to thank these projects for supporting this article.

#### References

- [1] W. Wu, "Ship communication network intrusion signal identification in accordance with hidden Markov model," *Journal of Coastal Research*, vol. 83, no. 1, pp. 868–871, 2018.
- [2] Zhao, A. Zhao, Yang, Peng, and Guo, "A modulation format identification method in accordance with information entropy analysis of received optical communication signal," *IEEE ACCESS*, vol. 23, no. 5, pp. 184–197, 2019.
- [3] Y. Liu and W. He, "Signal detection and identification in an optical camera communication system in moving state," *Journal of Physics: Conference Series*, vol. 13, no. 1, pp. 12282–12296, 2021.
- [4] G. J. Mendis, J. Wei, and A. Madanayake, "Deep learning based radio-signal identification with hardware design," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 11, no. 1, pp. 95–103, 2019.
- [5] J. Han, T. Zhang, D. Ren, and X. Zheng, "Communication emitter identification based on distribution of bispectrum amplitude and phase," *IET Science, Measurement & Technology*, vol. 11, no. 8, pp. 1104–1112, 2017.
- [6] S. Madbushi, R. Raut, R. Raut, and M. Rukmini, "A novel chaotic communication based test signal approach for identification of primary user emulation attack in cognitive radio networks," *International Journal of Intelligent Engineering and Systems*, vol. 11, no. 2, pp. 57–67, 2018.
- [7] Z. Zhao, A. Yang, G. Peng, and W. Tang, "A modulation format identification method in accordance with amplitude deviation analysis of received optical communication signal," *IEEE Photonics Journal*, vol. 4, no. 9, pp. 1–7, 2019.
- [8] J. Burek, M. Plodzien, L. Zylka, and P. Sulkowicz, "High-performance end milling of aluminum alloy: influence of different serrated cutting edge tool shapes on the cutting force," *Advances in Production Engineering & Management*, vol. 14, no. 4, pp. 494–506, 2019.

## Retraction

# Retracted: Managing the Gate Assignment Problem in the Hub Airport with Satellite Halls: A Transfer Demand-Oriented Approach

### Computational Intelligence and Neuroscience

Received 1 August 2023; Accepted 1 August 2023; Published 2 August 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] R. Wu, W. Zhou, L. Zhong, and Y. Liu, "Managing the Gate Assignment Problem in the Hub Airport with Satellite Halls: A Transfer Demand-Oriented Approach," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 5087941, 15 pages, 2022.



## Research Article

# Managing the Gate Assignment Problem in the Hub Airport with Satellite Halls: A Transfer Demand-Oriented Approach

Runfa Wu , Wenliang Zhou , Linhuan Zhong , and Yihan Liu 

*School of Traffic and Transportation Engineering, Central South University, Changsha, Hunan 410075, China*

Correspondence should be addressed to Wenliang Zhou; [zwl\\_0631@csu.edu.cn](mailto:zwl_0631@csu.edu.cn)

Received 6 January 2022; Accepted 15 March 2022; Published 25 April 2022

Academic Editor: Daqing Gong

Copyright © 2022 Runfa Wu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This study focuses on managing the gate assignment in the hub airport with both main terminal and satellite halls. We first formulate the gate assignment problem (GAP) as a binary linear programming model with multi-objective functions, where the practical constraints, e.g., gate time conflict and gate compatibility, are considered. Then, we incorporate the impact of gate assignment on transfer passengers and formulate the transfer demand-oriented gate assignment problem (TGAP) as a nonlinear model. A linearization approach and a heuristic approach are designed to solve the TGAP model. A case study is conducted based on the practical data of the Shanghai Pudong International Airport, where a comparison between the results of GAP and TGAP by the two proposed approaches is demonstrated. It shows that the proposed TGAP model and solution approaches can not only enhance the service for transfer passengers but also improve the gate utilization efficiency in the hub airport.

## 1. Introduction

Global air passenger demand and airport construction have experienced rapid growth over the past few decades, and passenger demand still grew steady in 2019 before the COVID-19 pandemic [1]. Driven by the recovery in domestic markets in China, many hub airports have returned to pre-pandemic levels in passenger demand and flight numbers; e.g., Shanghai Pudong International Airport was handling 110,000 domestic travelers and 900 domestic flights every day in March 2021, which are both more than that of 2019 [2]. As the import part of the hub-and-spoke system in air transportation, hub airports are experiencing rapid growth in passenger demand and flight numbers. To release operating pressure and provide better airplane/flight service and passenger service, many hub airports have built satellite halls connected with terminals by underground walkways or mass rapid transit (MRT) systems. This raises new challenges for efficient operations and passenger services in the hub airport.

Gates are a scarce resource at hub airports, facing intense air traffic and passenger demand pressure [3]. The gate operation connects air traffic (timetable and airplane

service), passenger service, and ground operations (including crew assignment), which makes it critical for efficient airport operations. The gate assignment problem (GAP) is to assign airplanes/flights to suitable boarding gates or the apron at the airport on an operating period (usually one day), according to the given flight timetable and airplane fleet assignment, also taking into account the airport layout, gate compatibility, airplane types, and so on. The typical objectives of GAP usually include two aspects, minimizing the operating costs and maximizing the efficiency of gate resources for airport operators; and maximizing satisfaction for passengers. The GAP has been extensively studied as one of the most important problems in the daily operations of the airport, and see [4, 5] for a detailed literature review, and we give an overview from airport and passenger perspectives.

From the airport operator perspective, the main objectives for GAP are efficient utilization of gate resources and reducing operating costs. Since the parking positions in the apron are usually far from the terminals and passengers need to take the shuttle bus. If airplanes are assigned to the apron, it will increase the waiting time of passengers and operation costs due to potential effects on ground operations and crew

assignment. The most common objective in GAP literature is to minimize the airplanes/flights assigned to the apron (Ding et al. [6], Dorndorf et al. [7, 8], Drexel and Nikulin [9], Deng et al. [10]). Some studies aim to minimize the total delay of airplanes when the airport is busy (Lim et al. [11], Kaliszewski et al. [12]). Moreover, as Karsu et al. [13] point out, hub airports may need to handle different types of flights (domestic/international) and airplanes, so minimizing the moves/costs of towing when airplanes need to move from a gate to another is also considered in the literature (Benlic et al. [14], Kumar and Bierlaire [15], Yu et al. [16]). We consider the gate compatibility instead of the towing operation in this study to capture the matching between airplane/flight types and gate facilities. It was also modeled in the objective function of Benlic et al. [14] and Neuman and Atkin [17]. Following the conclusion that the airport controls gate use will ensure they are used most efficiently in the survey of Gillen and Lall [18], we also consider the integrated operating of gates in terminal and satellite halls in the hub airport and include minimizing the number of used gates in the objective function.

Gate assignment also affects the passenger service quality in the airport [11], and it may influence the walking distance, waiting time, and transfer service of passengers. The consideration of passengers in the literature is mainly reflected in the objective function. Hub airports in metropolitan usually have multiterminals, and it may take a lot of time/distance to get the specific gate or transfer between gates, and many studies contribute to GAP to minimize walking time/distance of passengers. Bohr [19] proposed binary linear programming to minimize the passenger walking distance and solved it by the primal-dual simplex algorithm. Karsu et al. [13] formulated a mixed-integer nonlinear programming model for GAP to minimize the total walking distance of all passengers and the number of airplanes assigned to the apron and then proposed exact and heuristic approaches to solve it. Interested readers can also refer to Drexel and Nikulin [9], Haghani and Chen [20], Dell'Orco et al. [21], and Mokhtarimousavi et al. [22]. Besides, Yan and Huo [23] and Yan and Tang [24] are focused on minimizing the total passenger waiting time in GAP. However, from the survey of Entwistle [25], over 60% of passengers plan to shop in the airport, which means minimizing the waiting time is not always the objective of passengers, at least for some of them. Daş [26] proposed a multi-objective model to increase the shopping revenues in the airport through gate assignment, by minimizing the total passenger walking distance and assigning passengers to gates near the shop. We consider the transfer time budget of passengers in this study and propose a more comprehensive way to measure service for transfer passengers. A comparison with some abovementioned major-related studies is shown in Table 1.

In this study, we focus on the impacts of gate assignment on the service of transfer passengers in the hub airport with satellite halls. We proposed a novel transfer demand-oriented objective function considering the transfer time budget, together with objective functions of the airport operator, to explore the trade-off between airport operations and passenger service. Besides, we propose two optimization

models, namely a binary linear programming model for gate assignment problem (GAP) and a nonlinear model for transfer demand-oriented gate assignment problem (TGAP). A linearization approach and a heuristic approach are designed to solve the TGAP model, and then, case studies are performed using the data of Shanghai Pudong International Airport.

The remainder of this study is organized as follows. In Section 2, we give a detailed description of GAP and TGAP. The corresponding mathematic models are formulated in Section 3. Then, Section 4 develops the linearization approach and heuristic approach to solve the model. Case studies are illustrated in Section 5, and Section 6 concludes the study.

## 2. Problem Statement and Assumptions

**2.1. Problem Statement.** We consider a hub airport with the main terminal and several satellite halls, illustrated in Figure 1. Gates are available in both the terminal and satellite halls and integrated assigned by the airport operator, denoted by  $g \in G$ . Passengers can travel between the terminal and satellite halls via the MRT system. As shown in Figure 1, the airport also has parking positions for airplanes at the apron, denoted by  $G'$ , in case an airplane could not be assigned to a gate in the terminal or satellite halls.

The research period in this study is denoted by  $[T_1, T_2]$  (i.e., one day or one week) and discretized into equal-length time intervals  $t$ , and let  $t \in T$ . Some airplanes land and take off at the airport in this period, they occupy the gates for passenger arrivals and departures, and the set of airplanes is denoted by  $I$ . The services that provided by airplanes to transport passengers between airports are called a *flight*. For a specific airport, one airplane serves two flights, and we assume that the related flights of each airplane and the timetable are given. In Figure 2, we show the relationship between airplanes and related flights it serves. An airplane  $i \in I$  is considered, and it serves two flights: one arriving flight with arrival time  $a_i$  and one departing flight with departure time  $d_i$ , and it needs to occupy a gate  $g \in G$  or a parking position  $g \in G'$  during the time period  $[a_i, d_i]$ . Besides, the type of airplanes (wide/narrow-body) and related flights (domestic/international) are also given. Since the correspondence between aircrafts and flights is given, it is possible to model by either airplane or flight, and we use the airplane for modeling in this study.

As for the utilization of gates, the first thing to consider is time conflicts. As shown in Figure 3, airplanes  $i \in I$  and  $i' \in I$  use the same gate consecutively, and the usage time periods of the two airplanes  $[a_i, d_i]$  and  $[a_{i'}, d_{i'}]$  should not overlap. Besides, buffer time  $\tau_b$  should also be satisfied between serving two airplanes for ground operations. Secondly, because gates in terminal and satellite halls may have different functions (such as check-in facilities and passport control), we consider the gate compatibility in this study. In particular, for airplanes, we consider gate compatibility for wide/narrow-body types; for related flights, the gate compatibility is associated with serving domestic/international flights.

TABLE 1: Comparisons of major-related studies.

Paper	Airport-oriented objective	Passenger-oriented objective	Transfer passenger	Gate compatibility	Solution method
Bihr [19]	—	Walking distance	No	—	Simplex algorithm
Ding et al. [6]	Unassigned airplane	Walking distance	Yes	—	Tabu search
Yan and Tang [24]	—	Waiting time	No	—	Heuristic approach
Drexel and Nikulin [9]	Unassigned airplane and flight-gate preferences	Walking distance	Yes	—	Pareto SA
Dell'orco et al. [21]	Unassigned airplane	Walking distance	Yes	—	Bee Colony Algorithms
Das [26]	Passengers assigned to specific gates	Walking distance	No	—	Two phase local search
Deng et al. [30]	Gate utilization	Walking distance	Yes	Airplane type	Quantum evolutionary algorithm
This paper	Unassigned airplane and gate utilization	Transfer pressure	Yes	Airplane and flight type	Linearization approach and SA-based algorithm

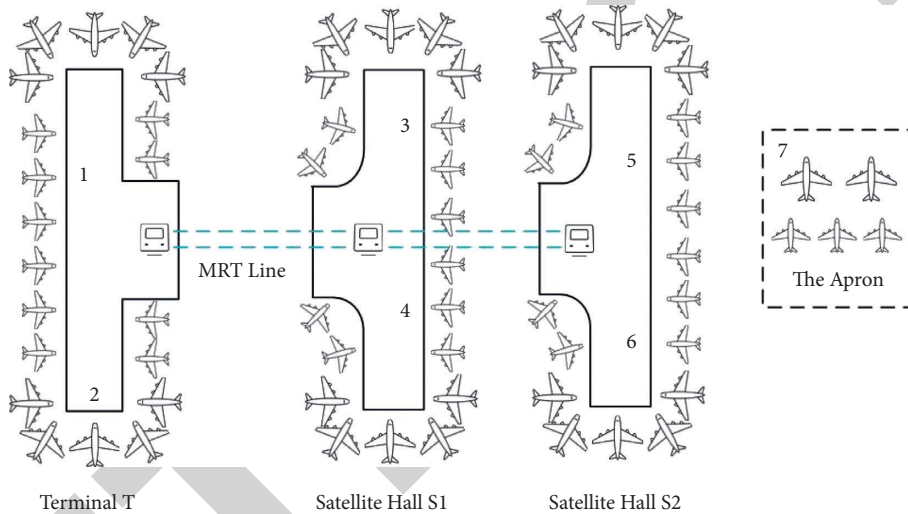


FIGURE 1: Example layout of the hub airport with satellite halls.

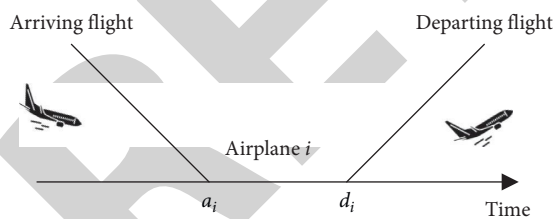


FIGURE 2: Airplane and served flights.

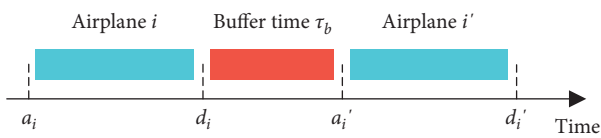


FIGURE 3: Gate utilization and buffer time.

The gate assignment problem (GAP) is to assign airplanes  $i \in I$  to gates or apron  $g \in G \cup G'$  with considering constraints such as gate time conflict constraints and gate

compatibility, and the objective function mainly concerns making efficient use of gates or reducing the number of occupied gates. The decision variables are binary gate assignment variables  $x_{ig}$ ,  $i \in I, g \in G$ , and binary gate utilization variables  $x_g$ ,  $g \in G$ . In particular,  $x_{ig}$  equals 1 if airplane  $i$  is assigned to gate  $g$  and otherwise equals 0;  $x_g$  equals 1 if gate  $g$  is used by any airplane and otherwise equals 0.

We also attempt to consider the impact of gate assignment on passengers in this study. The set of passenger groups is denoted by  $P$ . As shown in Figure 4, each group of passengers  $p \in P$  transfers from the same arrival flight served by airplane  $i_p^1 \in I$  to the same departure flight served by airplane  $i_p^2 \in I$  (which can be abbreviated as  $i_1$  and  $i_2$ ). The transfer time budget for passenger group  $p$  is defined as  $B_p = d_{i_2} - a_{i_1}$ . The number of passengers in group  $p$  is given and denoted by  $n_p$ .

Gate assignment determines passengers' shortest transfer time, including processing time, walking time, and MRT time. It will affect passengers' transfer in the airport, especially hub airports with satellite halls. The layout in

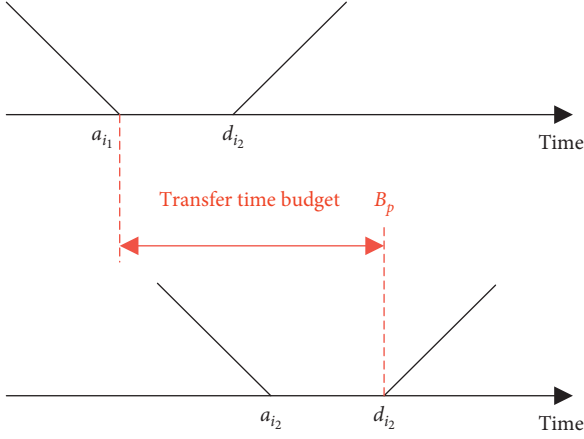


FIGURE 4: Illustration for transfer time budget of passengers.

Figure 1 is taken as an example, and a passenger whose transfer time budget is 60 min is considered. Different transfer scenarios are illustrated in Table 2 when arriving and departing airplanes are assigned to different gates. In Scenario 1, both gates of arriving and departing airplanes are assigned in terminal  $T$ , and the shortest transfer time is much less than the transfer time budget  $B_p$  and transfer success. Passenger transfer is more stressful in Scenario 2 because the assigned gates of two airplanes are located at  $T$  and  $S1$ , respectively. Passengers in Scenario 3 and Scenario 4 fail to transfer due to the gate assignment.

Thus, passengers  $p$  may fail to transfer if the shortest transfer time is too long between two gates assigned to airplanes  $i_1$  and  $i_2$ , i.e., exceeding the transfer time budget  $B_p$ . The demand-oriented gate assignment problem (TGAP) in this study focuses on the trade-off between transfer passenger service and gate utilization efficiency in GAP when satellite halls are constructed in the hub airport.

**2.2. Assumptions.** To facilitate the presentation of our studied problem in this study, the following assumptions are made:

**A1:** (airport layout). Considering a hub airport with one terminal, several satellite halls, and an apron, the shortest transfer time between any two gates is given. There is no limit on the number and type of airplanes that use the apron simultaneously.

**A2:** (flight and airplane). Given the flight timetable in the research period, including arrival/departure time, flight types (domestic/international), and airplane types (wide/narrow-body).

**A3:** (gate service). Only one airplane can use a gate at a time. All of the gates have the same buffer time  $\tau_b$  between serving two airplanes. The gate compatibility for flights and airplanes is considered.

**A4:** (passenger demand). Since the satellite halls mainly affect transfer passengers, it is assumed that we only consider the transfer passenger demand. The quantity, associated flights, and transfer time budget of passengers are given.

### 3. Mathematical Formulation

In this section, we first formulate the model for GAP to clarify the resource utilization and constraints in the hub airport, and then, we propose the model for TGAP considering the service of transfer passengers in Section 3.2.

**3.1. Notations and Decision Variables.** Table 3 lists general indices, sets, parameters, and variables in optimization models that appeared in this study.

**3.2. Model for GAP.** In this subsection, the mathematic model of GAP is formulated to integrate using the gates in terminal and satellite halls, including constraints and multi-objective functions.

**3.2.1. Constraints.** The constraints of GAP usually include gate utilization and airplane service, which are next described in detail.

(1) *Gate Time Conflict Constraint.* A feasible gate assignment scheme should guarantee that airplanes assigned to the same gate do not overlap in time and observe the buffer time. The airplane time incidence parameter  $\delta_{it}$  is introduced, which equals 1 when  $a_i \leq t \leq d_i + \tau_b$  and otherwise 0. So, we have the following:

$$\sum_{i \in I} \delta_{it} \cdot x_{ig} \leq 1, \quad \forall t \in T, \forall g \in G. \quad (1)$$

The incidence parameter  $\delta_{it}$  and assignment variable  $x_{ig}$  associate airplanes, gates, and time.

(2) *Gate Utilization Constraints.* For gate  $g \in G$ , if it is used by any airplane, the variable  $x_g$  equals 1, otherwise 0. So, we have gate utilization constraints that indicate the relationship between  $x_{ig}$  and  $x_g$  as follows:

$$\sum_{i \in I} x_{ig} \leq M \cdot x_g, \quad \forall g \in G, \quad (2)$$

where  $M$  is a sufficiently large positive constant.

(3) *Airplane Service Constraints.* Each airplane must and can only be assigned to one gate or the apron, and then:

$$\sum_{g \in G \cup G'} x_{ig} = 1, \quad \forall i \in I. \quad (3)$$

(4) *Gate Compatibility Constraints.* We consider the gate compatibility in this model, because gates in the different areas of terminal and satellite halls may have different functions, which are mainly influenced by facilities and equipment. The airplane gate compatibility incidence parameter  $\sigma_{ig}$  is introduced, which equals 1 if airplane  $i \in I$  can be served by gate  $g \in G$  and otherwise 0. We can derive the values of  $\sigma_{ig}$  based on the given airplane and gate types.

$$x_{ig} \leq \sigma_{ig}, \quad \forall i \in I, \forall g \in G \cup G'. \quad (4)$$

TABLE 2: Transfer scenarios for passenger group  $p$  under different gate assignment schemes.

Scenarios	Gate of arriving airplane	Gate of departing airplane	Shortest transfer time	Transfer time budget $B_p$	Result
1	Gate 1 (T)	Gate 2 (T)	30 min	60 min	Success
2	Gate 1 (T)	Gate 3 (S1)	55 min		Success
3	Gate 1 (T)	Gate 5 (S2)	70 min		Fail
4	Gate 1 (T)	Gate 7 (apron)	> 70 min		Fail

TABLE 3: Main sets, indices, parameters, and variables.

Symbol	Definition
<i>Sets and indices</i>	
$G$	Set of gates in the terminal and satellite halls
$G'$	Set of parking positions in the apron
$I$	Set of airplanes
$P$	Set of passenger groups
$T$	Set of discretized time intervals
$g$	Index of all gates and parking positions, $g \in G \cup G'$
$i$	Index of airplanes, $i \in I$
$p$	Index of passenger groups, $p \in P$
$t$	Index of discretized time intervals, $t \in T$
<i>Parameters</i>	
$[T_1, T_2]$	Daily operation period
$\tau_b$	Buffer time of gates
$a_i$	Arrival time of airplane $i$
$d_i$	Departure time of airplane $i$
$\delta_{it}$	Airplane time incidence parameter
$\sigma_{ig}$	Airplane gate compatibility incidence parameter
$i_p^1$	Arriving airplane of passenger group $p$ , abbreviated as $i_1$
$i_p^2$	Departing airplane of passenger group $p$ , abbreviated as $i_2$
$B_p$	Transfer time budget of passenger group $p$
$n_p$	Number of passengers in group $p$
$\tau(g_1, g_2)$	Shortest transfer time between gate $g_1 \in G \cup G'$ and $g_2 \in G \cup G'$
$\tau_c$	Waiting time for passengers when fail to transfer
$M$	A sufficiently large constant
$\varepsilon$	A sufficiently small positive constant
$c$	Weights in the objective function
<i>Variables</i>	
$x_{ig}$	Binary assignment variable: $x_{ig} = 1$ if airplane $i \in I$ is assigned to gate $g \in G \cup G'$ , and $x_{ig} = 0$ , otherwise
$x_g$	Binary utilization variable: $x_g = 1$ if gate $g \in G$ is used, and $x_g = 0$ , otherwise
$\varphi_p$	Transfer pressure of passenger group $p$
$y_{i_1, i_2, g_1, g_2}$	Binary variable: $y_{i_1, i_2, g_1, g_2} = 1$ for passenger $p$ , if arriving airplane $i_1$ is assigned to $g_1$ and departing airplane $i_2$ is assigned to $g_2$ , and $y_{i_1, i_2, g_1, g_2} = 0$ , otherwise

In particular, the apron can serve all types of airplanes and  $\sigma_{ig} = 1$ ,  $g \in G'$ .

##### (5) Constraints for Decision Variables

$$x_{ig} = \begin{cases} 1 & \text{if airplane } i \text{ is assigned to gate } g, \\ 0 & \text{otherwise,} \end{cases} \quad \forall i \in I, \forall g \in G \cup G',$$

$$x_g = \begin{cases} 1 & \text{if gate } g \text{ is used,} \\ 0 & \text{otherwise,} \end{cases} \quad \forall g \in G.$$
(5)

Constraints (5) and (6) are binary requirements on the decision variables.

**3.2.2. Objective Function.** Gates are the scarce resource in an airport, and operational efficiency depends on the utilization of this bottleneck resource. Since the hub airport

has terminal and satellite halls simultaneously, the GAP aims to make efficient use of the gates in the terminal and satellite halls and minimizes the operating costs.

An airplane can use parking positions in the apron if it cannot be assigned to a gate, but the apron is usually far away from the terminal and satellite halls, and passengers need to take a shuttle bus between the terminal and the apron. This will increase the transfer time of passengers on the one hand and increase the operating costs of the airport on the other hand. To make efficient use of the gates and avoid assigning airplanes to the apron, the first objective is to minimize the number of airplanes assigned to the apron:

$$\min Z_1 = \sum_{i \in I} \sum_{g \in G'} x_{ig}. \quad (6)$$

This is a common objective in the literature of GAP (Ding et al. [6], Dorndorf et al. [7, 8], Drexler and Nikulin [9], Deng et al. [10]). It is equivalent to maximizing the number of airplanes assigned to gates. Furthermore, this objective can be easily extended to maximize the total usage time of gates since the dwell time of each airplane is given, but it has no significant impact on passenger service so we use the objective (6) in this study.

Besides, the GAP is multi-objective in nature and operation costs for gates are expensive (including ground operation costs), which motivates us to consider objectives more comprehensively. Apart from minimizing the apron operation, the GAP also needs to consider the objective that minimizes the number of used gates, i.e.,

$$\min Z_2 = \sum_{g \in G} x_g. \quad (7)$$

In this study, we consider the objective function that consists of the number of airplanes assigned to the apron and the number of used gates and then formulate the GAP as a multi-objective optimization problem. For GAP in a hub airport with the apron, if we only consider the objective (6), gate operating hours may not be fully utilized in some situations; only consider the objective (7) is obviously not feasible either, which would assign all of the airplanes to the apron. Thus, the combination and trade-off between objectives (6) and (7) is comprehensive for GAP. We use linear weight to handle these two objectives and formulate the objective function as follows:

$$\min Z_{\text{GAP}} = \alpha_1 \cdot Z_1 + \alpha_2 \cdot Z_2, \quad (8)$$

where  $\alpha_1$  and  $\alpha_2$  are positive weights to denote the trade-off between objectives. In particular, we can obtain a Pareto-optimal solution if  $\alpha_1$  and  $\alpha_2$  are set 1, or their values are set according to the preference of airport operators.

**3.2.3. Mathematical Model for GAP.** The GAP can be formulated as follows:

$$\begin{cases} \min Z_{\text{GAP}} = \alpha_1 \cdot \sum_{i \in I} \sum_{g \in G'} x_{ig} + \alpha_2 \cdot \sum_{g \in G} x_g, \\ \text{s.t. constraints (1) - (6).} \end{cases} \quad (9)$$

The mathematical model for GAP given in (9) is binary integer linear programming. The model for GAP focuses on the gate resource optimization when a hub airport has both main terminal and satellite halls, and next, we will consider the impact of GAP on passengers and further improve the model involving different stakeholders.

**3.3. Model for TGAP.** Gate assignment affects the service quality of passengers, especially the transfer passengers in the hub airport where both main terminal and satellite halls are providing passenger service. Passengers may take a longer time to get from the arriving flight to the departing flight gate due to the gate assignment and may even exceed the transfer time budget resulting in a failed transfer. Thus, we will incorporate the service of transfer passengers in the GAP in this section.

With the given flight timetable and transfer scheme (arriving airplane  $i_1$  and departing airplane  $i_2$ ) of passenger group  $p \in P$ , we can get the transfer time budget  $B_p = d_{i_2} - a_{i_1}$ . The gates serving airplanes  $i_1$  and  $i_2$  are denoted as  $g_1$  and  $g_2$ , respectively. Given the layout of terminal and satellite halls in the airport, the shortest transfer time  $\tau(g_1, g_2)$  (including processing time, walking time, and MRT time) between any two gates is also fixed.

The gate assignment will influence the shortest transfer time of passenger group  $p$ . Here, we introduce the transfer pressure to describe the airport's service level for transfer passengers. The transfer pressure is the ratio of shortest transfer time to transfer time budget, and the transfer pressure for passenger  $p$  is denoted by  $\varphi_p$  and defined as follows:

$$\varphi_p = \begin{cases} \frac{\tau(g_1, g_2)}{B_p}, & \text{if } i_1, i_2 \text{ are both assigned to gates,} \\ \frac{\tau_c}{B_p}, & \text{otherwise,} \end{cases} \quad (10)$$

where  $\tau_c$  is the total transfer time (including process time and shuttle bus time) for passengers associated with the apron. If both associate airplanes of a passenger group are assigned to gates in the terminal and satellite halls, passengers need to do a gate-gate transfer with transfer time  $\tau(g_1, g_2)$ ; otherwise, passengers need to do a gate-apron, apron-gate, or even apron-apron transfer with transfer time  $\tau_c$ . Note that  $\tau_c$  is longer than the shortest transfer time between any two gates because the parking positions in the apron are usually far away from the terminal and satellite halls.

Then, we introduce the objective that minimizes the transfer pressure for passengers to capture the passenger service in the GAP of the hub airport; i.e.,

$$\min Z_3 = \sum_{p \in P} \sum_{g_1 \in G \cup G'} \sum_{g_2 \in G \cup G'} x_{i_1 g_1} x_{i_2 g_2} \varphi_p n_p. \quad (11)$$

The GAP considering passenger transfer time budget can be formulated as follows:

$$\begin{cases} \min Z_{\text{TGAP}} = \alpha_1 \cdot \sum_{i \in I} \sum_{g \in G'} x_{ig} + \alpha_2 \cdot \sum_{g \in G} x_g + \alpha_3 \cdot \sum_{p \in P} \sum_{g_1 \in G \cup G'} \sum_{g_2 \in G \cup G'} x_{i_1 g_1} x_{i_2 g_2} \varphi_p n_p, \\ \text{s.t. constraints (1) - (6).} \end{cases} \quad (12)$$



The mathematical model given in (12) explicitly considers the transfer passenger service and shows the trade-off between passenger service and operating costs of GAP in the hub airport. The mathematical model (12) is nonlinear programming, and the nonlinearities come from the objective function associated with passenger, where the calculation of transfer pressure  $\varphi_p$  is a segmentation function.

Table 4 presents the complexity of the model for GAP and TGAP. It can be seen that the model size depends on the number of gates, airplanes, passenger groups, and demand discretization (number of discretized time intervals). Suppose that there is a hub airport with 10 gates and 100 airplanes with 100 transfer passenger groups, the research period is [0: 00 – 24: 00]. If the discretization time interval is 5 min, there will be 120 variables and 2500 constraints in the GAP model (9). The number of variables in the TGAP model (11) is 1120, with the addition of variables related to transfer passengers.

#### 4. Solution Approach

The mathematic model (9) for GAP is a binary integer linear programming and can be solved by several existing commercial solvers, such as CPLEX and Gurobi (see, e.g., Linderoth and Ralphs [27]; Atamturk and Savelsbergh [28]).

As for the mathematic model (12) for TGAP, it is nonlinear programming with linear constraints, and we next propose two approaches to solve it.

**4.1. Linearization Approach.** In this section, the origin nonlinear programming model (12) will be transformed into binary integer linear programming by introducing new binary variables and linear constraints.

Focusing on the nonlinear objective function of the model (12), the calculation of transfer pressure  $\varphi_p$  is a segmentation function as shown in Eq. (11). According to the analysis of transfer time under different scenarios in Section 3.3, only need to set the shortest transfer time associated with  $g \in G'$  as  $\tau_c$ , i.e.,  $\tau(g_1, g_2) = \tau_c, g_1 \in G'$  or  $g_2 \in G'$  the objective that minimizes the transfer pressure in (12) could be updated as follows:

$$\min Z_3 = \sum_{p \in P} \sum_{g_1 \in G \cup G'} \sum_{g_2 \in G \cup G'} \frac{x_{i_1 g_1} x_{i_2 g_2} \tau(g_1, g_2) n_p}{B_p}. \quad (13)$$

It can be observed that (13) is nonlinear because of productions of binary variables  $x_{i_1 g_1}$  and  $x_{i_2 g_2}$ , and they can be replaced by auxiliary binary variables  $y_{i_1, 2, g_1, 2}$ . Following Williams [29], the productions can be replaced by adding linear constraints:

$$\begin{cases} -x_{i_1 g_1} + y_{i_1, 2, g_1, 2} \leq 0 \\ -x_{i_2 g_2} + y_{i_1, 2, g_1, 2} \leq 0 \\ x_{i_1 g_1} + x_{i_2 g_2} - y_{i_1, 2, g_1, 2} \leq 1 \\ y_{i_1, 2, g_1, 2} \in \{0, 1\} \end{cases}, \quad \forall p \in P, \forall g_1, g_2 \in G. \quad (14)$$

Thus, the linearized model for TGAP considering the transfer passenger service can be formulated as follows:

$$\begin{cases} \min Z_{\text{TGAP}} = \alpha_1 \cdot \sum_{i \in I} \sum_{g \in G'} x_{ig} + \alpha_2 \cdot \sum_{g \in G} x_g + \alpha_3 \cdot \sum_{p \in P} \sum_{g_1 \in G \cup G'} \sum_{g_2 \in G \cup G'} \frac{y_{i_1, 2, g_1, 2} \tau(g_1, g_2) n_p}{B_p}, \\ \text{s.t. constraints (1) – (6), (14)}. \end{cases} \quad (15)$$

Note that model (15) is linear programming and can easily be solved by commercial solvers such as CPLEX and Gurobi to find a globally optimal solution.

As shown in Table 4, the number of auxiliary binary variables  $y_{i_1, 2, g_1, 2}$  is  $|P| \times (|G| + 1)^2$ . Based on the example in Section 3, there will be 121120 variables and 489000 constraints in the linearized model (15) for TGAP. Besides, when the number of passenger groups and airport gates increases, the number of variables and constraints increases rapidly, which takes a long computation time to solve the TGAP with commercial solvers. To address this issue, we further design a heuristic approach to solve the TGAP.

**4.2. Heuristic Approach.** The gate assignment problem is a complex nondeterministic polynomial hard (NP-hard) problem due to the complex layout of airports, multi-flights, passenger trips, and gate compatibility [30, 31], and many studies adopted heuristic approaches to solve it [16, 21, 24]. To solve TGAP in large hub airports requires an efficient

algorithm to obtain a satisfactory solution and solve the problem in reasonable CPU time. The simulated annealing (SA) algorithm is a metaheuristic to approximate the global optimization and has good robustness. Thus, we propose an approach for TGAP at large hub airports based on the framework of the SA algorithm.

In Algorithm 1, we adopt the following strategies to adjust the assignment scheme and get neighborhood solutions.

**4.2.1. Initial Solution.** The model (11) for TGAP has the same constraints as the model (9) for GAP, and we can use the optimal solution of model (9) as the initial solution. It is already efficient in terms of gate resource utilization. As for the model scales, note that the model (9) can be decomposed into two subproblems by gate compatibility on airplane types: one assignment for wide-body airplanes and associate gates and another for narrow-body airplanes and associate gates. In this way, the initial solution of TGAP is designed.



TABLE 4: Numbers of variables and constraints in the models.

Variables and constraints	Quantity
<i>Variables</i>	
$x_{ig}$	$ I  \times  G  +  I $
$x_g$	$ G $
$\varphi_p$	$ P $
$y_{i_1,2} \theta_{i_2}$	$ P  \times ( G  + 1)^2$
<i>Constraints</i>	
Gate time conflict constraint (1)	$ T  \times  G $
Gate utilization constraints (2)	$ G $
Airplane service constraints (3)	$ I $
Gate compatibility constraints (4)	$ I  \times  G  +  I $
Binary constraints (5)	$ I  \times  G  +  I $
Binary constraints (6)	$ G $

Note:  $|\cdot|$  represents the cardinality of a set.

**4.2.2. Passenger Service Adjustment Strategy.** The passenger service adjustment strategy aims to reduce the transfer pressure of passengers, which includes three options: insert option, swap operation, and remove operation. These operations are performed sequentially, with only one of them executed in each loop, and detailed options are shown as follows.

In the current solution, we already know the gate assignment scheme, i.e., the specific gate of each airplane. As we know the transfer information of passenger group  $p \in P$ ,  $P_i^1$  and  $P_i^2$  are denoted as the subset of passenger groups whose arriving and departing airplane is  $i \in I$ , respectively. Then, we denote  $\varphi_i$  as the total transfer pressure of passengers associated with airplane  $i$ , which is given as follows:

$$\varphi_i = \sum_{p \in P_i^1} \varphi_p n_p + \sum_{p \in P_i^2} \varphi_p n_p. \quad (16)$$

The total transfer pressure of each airplane is calculated by equation (16) and does the following operations: ① the airplane that has the maximum transfer pressure and is already assigned to a gate is selected. ② Insert operation: the subset of gates that has available time and already been used is found, the selected airplane is inserted into one of them randomly. If the subset is empty, the next operation is proceeded. ③ Swap operation: the subset of airplanes that has the same time interval and type (no violation of constraints (1) and (4)) is found, the selected airplane and one of them are randomly swapped. If the subset is empty, the next operation is proceeded. ④ Remove operation: if the above two operations are not executed, the subset of gates suitable for the selected airplane is found, one of them is chosen at random, and the selected airplane is assigned to this gate, and then, the airplanes with conflicts are assigned to the apron.

**4.2.3. Gate Utilization Strategy.** The gate utilization strategy concentrates on reducing the number of used gates, which includes two options: insert option and remove operation. Based on the current solution, we could know the set of airplanes assigned to gate  $g \in G$  and denoted by  $I_g$ .  $u_g$  is denoted as the time utilization ratio of gate  $g$ , which equals the ratio of occupied time to research period:

$$u_g = \sum_{i \in I_g} \frac{(\tau_b + d_i - a_i)}{(T_2 - T_1)}. \quad (17)$$

To reduce the number of used gates, the following operations are executed: ① the gate with the lowest time utilization ratio using (17) is found, and if the ratio is lower than a threshold value (such as 40%), then next options are done. ② insert operation: for airplanes assigned to the selected gate, the subset of gates that has available time and already used is found, and these airplanes are inserted in one of them randomly. ③ Remove operation: the rest of the airplanes are assigned to the apron after the previous option.

**4.2.4. Apron Airplane Adjustment Strategy.** The apron airplane adjustment strategy is designed to reduce the number of airplanes assigned to the apron and transfer pressure of passengers. Thus, included operations in this part of the strategy consider both of the two objectives. ① The total transfer pressure of each airplane in the apron by (16) is calculated and the maximum one of them is found. ② The insert and swap operations are executed. This airplane is randomly inserted into a gate that has available time periods; otherwise, it is swapped with another airplane assigned to a gate without violation of constraints (1) and (4). ③ If the above two operations are not satisfied, one of the airplanes in the apron is attempted insert into an available gate. ④ If there is any available empty gate, the airplane with the shortest overlap time period with other airplanes in the apron is found and assigned to the new gate.

The proposed heuristic approach shown in Algorithm 1 is based on the framework of the SA algorithm, together with three strategies for improving different parts of objective functions, and it would find a satisfactory solution of TGAP in the model (13).

## 5. Case Studies

To demonstrate the performance of the proposed models and solve approaches, we use the data of the Shanghai Pudong International Airport in China as a case study. We will describe the experiment data in Section 5.1. The numerical results are presented in Section 5.2.

All numerical tests are conducted on a personal computer with Intel® Core (™) 3.00 GHz processor and 16.00 GB RAM and Windows 10 Home Edition Operating System (64 bit). The YALMIP-R20190425 together with MATLAB R2019b is used to conduct the numerical tests. The commercial solver Gurobi optimization studio 8.1.1 (with academic license) is adopted to solve GAP and linearized TGAP models, and the solver used the branch-and-cut algorithm to find optimal solutions for the above two mixed-integer programming models.

**5.1. Data and Parameter Setting.** We consider a real-world case study on the Shanghai Pudong International Airport, which is an important hub airport in eastern China. The gates are integrated used in a terminal  $T$  and connected

```

(1) Input Solution  $S_0$  of GAP as the initial solution of TGAP, passenger information (flights and volume), algorithm parameters (initial and end temperature  $T_0, T_f$ , the maximum number of inner iterations  $m_{\max}$  and temperature drop ratio  $\theta$ ).
(2) Calculate the objective value  $Z_0$  in the model (12). Set current solution  $S \leftarrow S_0$ ,  $Z \leftarrow Z_0$  and the best solution  $S^* \leftarrow S_0$ ,  $Z^* \leftarrow Z_0$ ; the current temperature  $T \leftarrow T_0$ .
(3) Repeat
(4)    $m \leftarrow 0$ ;
(5)   Repeat
(6)      $m \leftarrow m + 1$ ;
(7)     Neighborhood searching using strategies for different objectives in Section 4.2
(8)       ① search for the neighborhood by passenger service strategy;
(9)       ② search for the neighborhood by gate utilization strategy;
(10)      ③ search for the neighborhood by apron airplane strategy.
(11)    Get the neighborhood solution  $S'$  and  $Z'$ .
(12)    Update the current solution and best solution
(13)       $\Delta Z \leftarrow Z' - Z$ ;
(14)      if ( $\Delta Z \leq 0$ ) then
(15)         $S \leftarrow S'$ ,  $Z \leftarrow Z'$  and  $S^* \leftarrow S'$ ,  $Z^* \leftarrow Z'$ ;
(16)      else if ( $(\exp[-\Delta Z/T]) > \text{Random}(0, 1)$ ) then
(17)         $S \leftarrow S'$ ,  $Z \leftarrow Z'$ ;
(18)      end
(19)    Until  $m = m_{\max}$ 
(20)     $T \leftarrow \theta \cdot T$ ;
(21)  Until  $T < T_f$ 
(22) Output optimal TGAP solution  $S^*$  and  $Z^*$ 

```

ALGORITHM 1: Solving framework based on SA algorithm for TGAP.

satellite hall S, and both terminal T and satellite hall S could handle transfer processes for passengers. There is an MRT line that connects T and S to quickly transport passengers, assuming that passengers' MRT time for a one-way trip is 5 minutes (for layout, see Figure 5). We consider the GAP and TGAP for 28 gates in terminal T and 41 gates in satellite hall S and an apron, and the detailed information of gates can be found in Table 5. The compatibility of gates such as service for domestic/international flights and wide/narrow-body airplanes is also given.  $\tau_c$  is set as 180 minutes.

In this case study, the considered research period is set as [0: 00 – 24: 00], which covers a full day of operations. We select 296 airplanes related to the above gates of China Eastern Airlines and Xiamen Airlines on January 20, 2018. Table 6 shows several records as an example, and every record corresponds to one airplane, which services two flights. The information of airplanes includes the arrival and departure date, arrival and departure time, arrival and departure flights, airplane types (wide/narrow-body), and flight types (domestic/international).

Meanwhile, transfer information of more than 3000 passengers is selected and divided into groups based on arrival and departure flights. The example of information is shown in Table 7, which includes arrival and departure flight, arrival and departure date, and passenger number in groups. Combining with the information of airplanes and flights in Table 6, we can easily get the transfer time budget of each passenger group. Since the layout of the airport is set, the shortest transfer time  $\tau(g_1, g_2)$  between any two gates is also determined, including processing time, walking time, and MRT time. The buffer time of gates  $\tau_b$  is set to 45 minutes.

The linear weights in the model (10), (13), and (16) are, respectively,  $\alpha_1 = 296$ ,  $\alpha_2 = 1$ ,  $\alpha_3 = 10$ ; the algorithm parameters are set as follows:  $T_0 = 10^6$ ,  $T_f = 5$ ,  $m_{\max} = 50$ ,  $\theta = 0.5$ .

**5.2. Computational Results.** Given the above data and settings, the proposed solution approaches will be implemented for GAP and TGAP. The results of GAP, TGAP with linearization approach, and performance comparison are shown in Subsection 5.2.1; Subsection 5.2.2 shows the result of TGAP solved by the heuristic approach.

**5.2.1. Solutions of GAP and Linearized TGAP.** In this subsection, we solve the GAP and linearized TGAP by commercial solver Gurobi optimization, and the CPU time to solve GAP is 5.75s. The result of GAP is shown in Figure 6, and the horizontal axis represents the research period ([0: 00 – 24: 00] on January 20, 2018), and the vertical axis represents the total of 69 gates in terminal T and satellite hall S. The colored bars in Figure 6 represent the time period when the airplanes occupy the corresponding gates, and the buff time is not included. It can be clearly seen that airplanes satisfy the time conflict constraints of gates and buffer time is also held between two adjacent airplanes. One can find that airplanes arriving and departing during [0: 00 – 6: 00] usually occupy the gates for a long time as passengers tend not to travel at this period. The result of TGAP is shown in Figure 7, which also satisfies all of the constraints.

From Figure 6, we can find that several gates (S29, S30, S39, and S41) are not occupied by any airplane, and they are all serving wide-body airplanes. Meanwhile, Figure 8

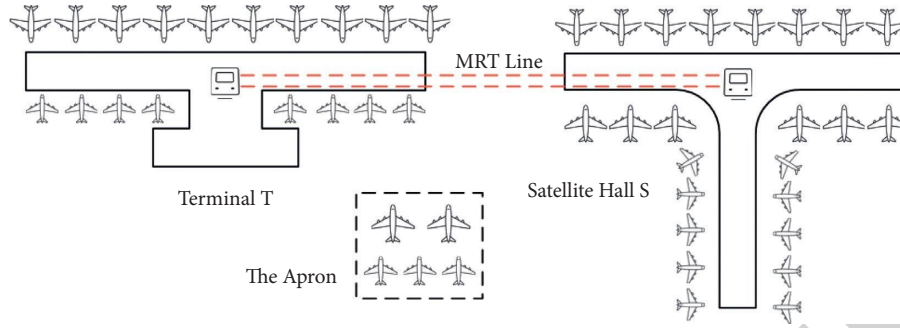


FIGURE 5: Layout of terminal T1 and satellite hall S1 in Shanghai Pudong International Airport.

TABLE 5: Gate information and compatibility in the case study.

Gate no.	Terminal	Arrival service	Departure service	Airplane type	Gate no.	Terminal	Arrival service	Departure service	Airplane type
T1	T	I	I	N	S1	S	D	D	N
T2	T	I	I	W	S2	S	D	D	N
T3	T	I	I	W	S3	S	D	D	N
T4	T	I	I	W	S4	S	D	D	N
T5	T	I	D, I	W	S5	S	D	D	N
T6	T	D, I	D, I	W	S6	S	D	D	N
T7	T	D, I	D, I	N	S7	S	D	D	N
T8	T	D, I	D	N	S8	S	D	D	N
T9	T	D, I	D	N	S9	S	D	D	N
T10	T	D	D	N	S10	S	D	D	N
T11	T	D	D	N	S11	S	I	I	N
T12	T	D	D	N	S12	S	I	I	N
T13	T	D	D	N	S13	S	I	I	N
T14	T	D	D	N	S14	S	D	D	N
T15	T	D	D	N	S15	S	D	D	N
T16	T	D	D	N	S16	S	D	D	N
T17	T	D	D	N	S17	S	D	D	N
T18	T	D	D	N	S18	S	D	D	N
T19	T	D	D	N	S19	S	D	D	N
T20	T	D	D, I	N	S20	S	D	D	N
T21	T	D	D, I	N	S21	S	D	D	N
T22	T	D, I	D, I	N	S22	S	D	D	N
T23	T	D, I	D, I	W	S23	S	D	D	N
T24	T	D, I	D, I	W	S24	S	D	D	N
T25	T	D, I	I	W	S25	S	D	D	N
T26	T	I	I	W	S26	S	D	D	N
T27	T	I	I	W	S27	S	D	D	N
T28	T	I	I	W	S28	S	D	D	N
S29	S	D	D	W	S36	S	I	I	W
S30	S	D	D	W	S37	S	I	I	W
S31	S	I	I	W	S38	S	I	I	W
S32	S	I	I	W	S39	S	I	I	W
S33	S	I	I	W	S40	S	I	I	W
S34	S	I	I	W	S41	S	I	I	W
S35	S	I	I	W					

Note. T: terminal, S: satellite hall, D: domestic flight service, I: international flight service, W: wide-body airplane service, N: narrow-body airplane service.

shows that all wide-body airplanes are already assigned to gates in GAP and that ratio for narrow-body airplanes is 81%, and the total number of airplanes successfully assigned to the gates is 249. Here is why we consider the objective that minimizes the number of used gates in (7), which can improve the efficiency of gate utilization when one kind of resource is sufficient, i.e., gates for wide-body

airplanes. Thus, the combination of objectives (6) and (7) is more comprehensive for gate utilization in both GAP and TGAP.

Next, we compare the solution of GAP and TGAP, and values of different parts in objective functions are shown in Table 8, of which the value of  $Z_3$  (transfer pressure) in GAP is calculated based on the optimal solution of the model (10)

TABLE 6: Partial airplane and flight records.

Airplane no.	Arrival flight number	Arrival date and time	Arrival type	Departure flight number	Departure date and time	Departure type	Airplane type
1	MU6753	Jan 19, 19:50	D	MU6358	Jan 20, 08:15	D	N
2	MU6785	Jan 20, 11:00	D	MU398	Jan 20, 13:10	I	N
3	MU6155	Jan 20, 13:20	I	MU6494	Jan 20, 14:25	D	N
4	MU663	Jan 20, 16:30	I	MU6588	Jan 20, 20:05	I	W
5	FM9188	Jan 20, 17:20	D	FM865	Jan 20, 18:20	I	N
6	MU5588	Jan 20, 21:20	D	MU5515	Jan 21, 08:55	D	N

Note. D: domestic flight, I: international flight, W: wide-body airplane, N: narrow-body airplane.

TABLE 7: Partial transfer passengers' information.

Passenger group no.	Number of passengers	Arrival flight number	Arrival date and time	Departure flight number	Departure date and time
1	10	MU5668	Jan 19, 18:35	MU739	Jan 20, 00:20
2	44	MU212	Jan 20, 10:20	MU207	Jan 20, 12:05
3	9	MU5698	Jan 20, 14:30	MU545	Jan 20, 16:10
4	20	MU504	Jan 20, 18:25	MU6720	Jan 20, 20:55
5	7	MU5130	Jan 20, 21:40	MU5023	Jan 21, 00:15

TABLE 8: Performance comparison of GAP and linearized TGAP solutions.

Problem	$Z_1$	$Z_2$	$Z_3$	Gate time utilization rate
GAP	47	65	1361.15*	66.40%
TGAP	48	65	1039.31	70.21%

\*Note: the total passenger transfer pressure  $Z_3$  of GAP is calculated after solving the optimal solution.

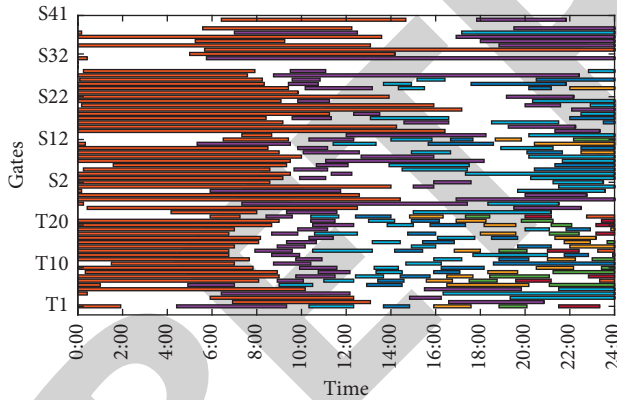


FIGURE 6: Optimal gate assignment scheme of GAP.

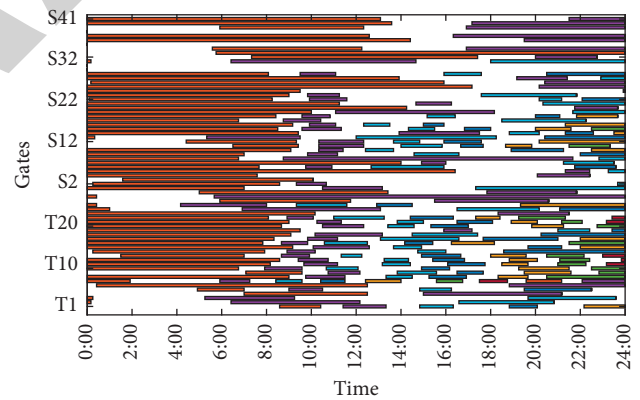


FIGURE 7: Optimal gate assignment scheme of linearized TGAP.

and services for passengers are not taken into account. It is obvious from Table 8 that when the first and second objectives (minimizing the number of apron airplanes  $Z_1$  and used gates  $Z_2$ ) are close in the GAP and TGAP, considering the objective  $Z_3$  can significantly reduce the transfer pressure of passengers (23.64% reduction). That is, model (12) of the TGAP can improve the service level of transfer passengers without increasing the resource requirement in the hub airport. Moreover, although the number of assigned airplanes has decreased in the TGAP solution, the time utilization rate of gates in terminal  $T$  and satellite hall  $S$  has increased, and the scheme in TGAP assigns airplanes serving more transfer passengers and with less occupation time to

the gates. This also demonstrates that considering the gate resource utilization and passenger service simultaneously is a more comprehensive way to address gate assignment in the hub airport.

5.2.2. *Solution of TGAP Adopting Heuristic Approach.* Although the proposed linearization approach in Section 4.1 can obtain the global optimal solution for TGAP, it takes a long time to converge. We adopted the commercial solver Gurobi optimization studio 8.1.1 to solve the linearized TGAP, and the CPU time to get the optimal solution is 6.09 h. The proposed heuristic approach takes

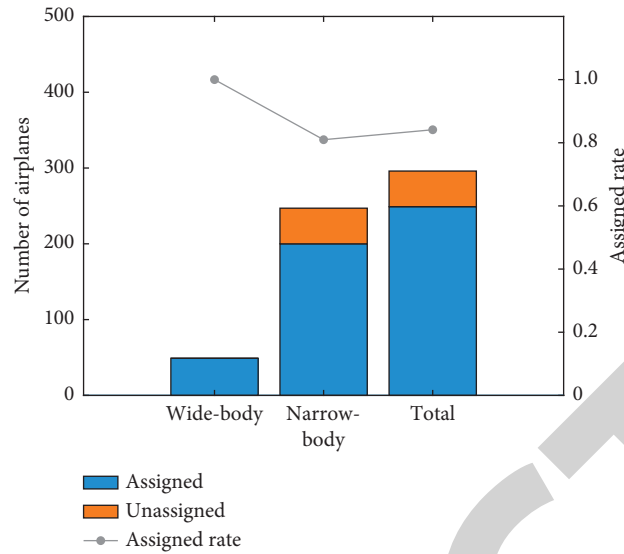


FIGURE 8: Assigned numbers and rates for different types of airplanes in GAP.

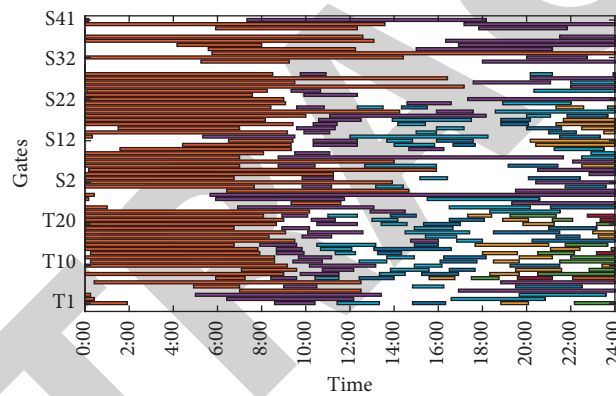


FIGURE 9: Gate assignment scheme of TGAP solved by the heuristic approach.

726s to solve the TGAP with the same input and parameters and only has a 2.59% solution GAP with linearization approach, and the gate assignment scheme is shown in Figure 9 as follows.

Passenger group no. 3 in Table 7 is taken as an example, passenger transfers from flight MU5698 arriving on January 20, 14:30, to flight MU545 departing at January 20, 16:10, and the transfer time budget is 100 min. In the GAP solution, airplanes separately serving the arrival and departure flight are assigned to gates S5 and T3. As the two gates are not in the same terminal, the shortest transfer time for passengers is 60 min. When we take the transfer pressure into account in TGAP, the above two airplanes are assigned to gates T18 and T5, which are both in the terminal *T*, passengers' shortest transfer time reduced to 35 min, and transfer pressure decreased from 0.60 to 0.35. Furthermore, the total transfer pressure of passengers in the solution of TGAP by the heuristic approach is still lower than that in the solution of GAP. This indicates that the TGAP well considered the service for passengers and realized the integrated assignment of gates in terminal and satellite halls.

As shown in Figure 10, gates in the terminal *T* are all used due to the shorter transfer time related to the gates in *T* than that in *S*, and there are 3 gates in the satellite hall *S* serving wide-body airplanes that are not used. The total utilization rate of gates is 96%, and the time utilization rate is 69.98%. For TGAP, although the gate time utilization rate in the solution of the linearized approach (70.21%) is higher than that of the heuristic approach, the difference is not significant.

Next, we compare the passenger transfer pressure in three cases: the GAP solution, linearized approach of TGAP solution, and heuristic approach of TGAP solution. The proportion of passengers within different transfer pressure intervals of the above three solutions are reported in Figure 11, where we observe that most of the passengers' transfer pressure remains at a relatively low level and lies in the range of [0.1, 0.5] in all three cases. Two solutions of TGAP are compared with the GAP solution, and we can see that the quantity of passengers who experience low transfer pressure ([0.1, 0.3]) in TGAP solutions is significantly more than that in the GAP solution, while in the high transfer

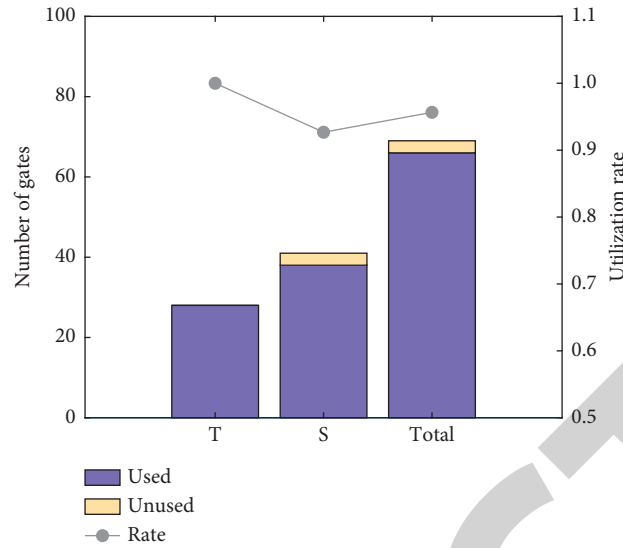


FIGURE 10: Gate utilization in terminal T and satellite hall S in TGAP.

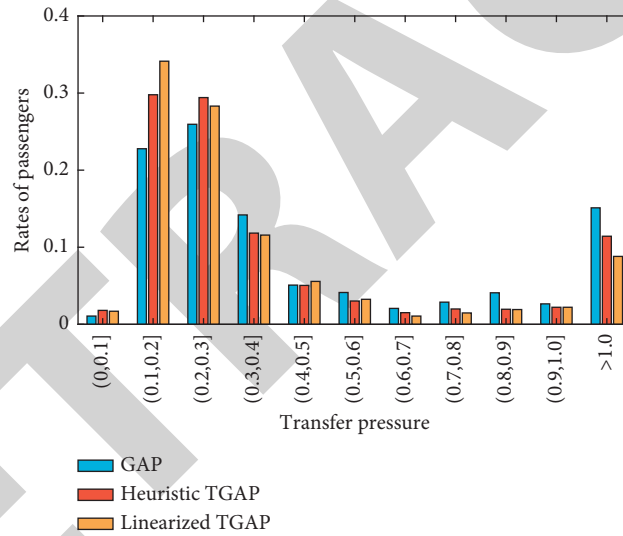


FIGURE 11: Comparison of passenger transfer pressure.

pressure range  $([0.5, 1.0])$ , the proportion of passengers in TGAP solutions is less than that in the GAP solution. This result indicates that proposed TGAP models could improve the service for transfer passengers.

Turning now to solutions of TGAP by linearization approach and heuristic approach, the distributions of passenger transfer pressure in these two solutions are comparable, which means the proposed heuristic approach could obtain a satisfactory solution in a reasonable time. What is striking in Figure 11 is that some passengers in all three solutions have transfer pressure greater than 1 because related airplanes are assigned to the apron, and this situation is enhanced in TGAP solutions.

## 6. Conclusions and Future Research

In this study, we focus on the impacts of gate assignment on the service of transfer passengers in the hub airport with satellite halls. First, a binary linear programming model for GAP is proposed that considers the gate time conflict, gate compatibility constraints, and the airport operator-oriented objective functions. Then, we introduce the transfer time budget and transfer pressure to measure the passenger service and formulate the TGAP as a nonlinear programming with linear constraints. In particular, multi-objective functions were considered in the TGAP model, including transfer demand-oriented and operator-oriented objectives.

We proposed a linearization approach and a SA algorithm-based heuristic approach to solve the nonlinear model of TGAP. Finally, the case study based on practical data demonstrated the benefits of the proposed models and solution approaches. In the experimental results, it was verified that the proposed TGAP model and solution approaches can improve the service for transfer passengers and lead to more efficient utilization of gate resources in the hub airport.

Further research could consider the randomness of transfer passenger demand and the effects of random flight delay on gate assignment and transfer passenger service. We can also manage the fairness of passengers through transfer pressure in the gate assignment problem.

## Data Availability

The practical data used to support the case study and findings in this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this study.

## Acknowledgments

This work was supported by research grants from the National Natural Science Foundation of China (grant nos. U1934216, 71871226, and U2034208) and the Fundamental Research Funds for the Central Universities of Central South University (grant no. 2019zzts272).

## References

- [1] I. A. T. A pressroom, "Slower but steady growth in 2019. International air transport association," 2020, <https://www.iata.org/en/pressroom/pr/2020-02-06-01/>.
- [2] Shanghai Municipal People's Government, "Daily flights nearly back to normal," Shanghai Municipal People's Government, 2021, <http://english.shanghai.gov.cn/nw48081/20210329/97de7bc7f58948d081010177cdeae2e6.html>.
- [3] M. A. Ornek, C. Ozturk, and I. Sugut, "Integer and constraint programming model formulations for flight-gate assignment problem," *Operational Research*, vol. 23, pp. 1–29, 2020.
- [4] U. Dorndorf, A. Drexl, Y. Nikulin, and E. Pesch, "Flight gate scheduling: state-of-the-art and recent developments," *Omega*, vol. 35, no. 3, pp. 326–334, 2007.
- [5] G. S. Daş, F. Gzara, and T. Stütze, "A review on airport gate assignment problems: single versus multi objective approaches," *Omega*, vol. 92, Article ID 102146, 2020.
- [6] H. Ding, A. Lim, B. Rodrigues, and Y. Zhu, "The over-constrained airport gate assignment problem," *Computers & Operations Research*, vol. 32, no. 7, pp. 1867–1880, 2005.
- [7] U. Dorndorf, F. Jaehn, and E. Pesch, "Modelling robust flight-gate scheduling as a clique partitioning problem," *Transportation Science*, vol. 42, no. 3, pp. 292–301, 2008.
- [8] U. Dorndorf, F. Jaehn, and E. Pesch, "Flight gate scheduling with respect to a reference schedule," *Annals of Operations Research*, vol. 194, no. 1, pp. 177–187, 2012.
- [9] A. Drexl and Y. Nikulin, "Multicriteria airport gate assignment and Pareto simulated annealing," *IIE Transactions*, vol. 40, no. 4, pp. 385–397, 2008.
- [10] W. Deng, H. Zhao, X. Yang, J. Xiong, M. Sun, and B. Li, "Study on an improved adaptive PSO algorithm for solving multi-objective gate assignment," *Applied Soft Computing*, vol. 59, pp. 288–302, 2017.
- [11] A. Lim, B. Rodrigues, and Y. Zhu, "Airport gate scheduling with time windows," *Artificial Intelligence Review*, vol. 24, no. 1, pp. 5–31, 2005.
- [12] I. Kaliszewski, J. Miroforidis, and J. Stańczak, "The airport gate assignment problem - multi-objective optimization versus evolutionary multi-objective optimization," *Computer Science*, vol. 18, no. 1, 41 pages, 2017.
- [13] Ö. Karsu, M. Azizoğlu, and K. Alanlı, "Exact and heuristic solution approaches for the airport gate assignment problem," *Omega*, vol. 103, Article ID 102422, 2021.
- [14] U. Benlic, E. K. Burke, and J. R. Woodward, "Breakout local search for the multi-objective gate allocation problem," *Computers & Operations Research*, vol. 78, pp. 80–93, 2017.
- [15] V. Prem Kumar and M. Bierlaire, "Multi-objective airport gate assignment problem in planning and operations," *Journal of Advanced Transportation*, vol. 48, no. 7, pp. 902–926, 2014.
- [16] C. Yu, D. Zhang, and H. Y. K. Lau, "An adaptive large neighborhood search heuristic for solving a robust gate assignment problem," *Expert Systems with Applications*, vol. 84, pp. 143–154, 2017.
- [17] U. M. Neuman and J. A. D. Atkin, "Airport gate assignment considering ground movement," in *Proceedings of the International Conference on Computational Logistics*, pp. 184–198, Springer, Berlin, Heidelberg, September 2013.
- [18] D. Gillen and A. Lall, "Developing measures of airport productivity and performance: an application of data envelopment analysis," *Transportation Research Part E: Logistics and Transportation Review*, vol. 33, no. 4, pp. 261–273, 1997.
- [19] R. A. Bihl, "A conceptual solution to the airplane gate assignment problem using 0, 1 linear programming," *Computers & Industrial Engineering*, vol. 19, no. 1-4, pp. 280–284, 1990.
- [20] A. Haghani and M.-C. Chen, "Optimizing gate assignments at airport terminals," *Transportation Research Part A: Policy and Practice*, vol. 32, no. 6, pp. 437–454, 1998.
- [21] M. Dell'Orco, M. Marinelli, and M. G. Altieri, "Solving the gate assignment problem through the fuzzy bee colony optimization," *Transportation Research Part C: Emerging Technologies*, vol. 80, pp. 424–438, 2017.
- [22] S. Mokhtarimousavi, D. Talebi, and H. Asgari, "A non-dominated sorting genetic algorithm approach for optimization of multi-objective airport gate assignment problem," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2672, no. 23, pp. 59–70, 2018.
- [23] S. Yan and C.-M. Huo, "Optimization of multiple objective gate assignments," *Transportation Research Part A: Policy and Practice*, vol. 35, no. 5, pp. 413–432, 2001.
- [24] S. Yan and C.-H. Tang, "A heuristic approach for airport gate assignments for stochastic flight delays," *European Journal of Operational Research*, vol. 180, no. 2, pp. 547–567, 2007.
- [25] M. Entwistle, "Customer service and airport retail: stimulate passenger spending," *Journal of Airport Management*, vol. 1, no. 2, pp. 151–157, 2007.
- [26] G. S. Daş, "New multi objective models for the gate assignment problem," *Computers & Industrial Engineering*, vol. 109, pp. 347–356, 2017.
- [27] J. T. Linderoth and T. K. Ralphs, "Noncommercial software for mixed-integer linear programming," in *Integer*



## *Retraction*

# **Retracted: Optimal Defect Detection and Sensing System of Railway Tunnel Radar considering Multisensor System Combined with Active Interference Suppression Algorithm**

### **Computational Intelligence and Neuroscience**

Received 17 October 2023; Accepted 17 October 2023; Published 18 October 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### **References**

- [1] Y. Lei, Y. Zou, B. Jiang, and T. Tian, "Optimal Defect Detection and Sensing System of Railway Tunnel Radar considering Multisensor System Combined with Active Interference Suppression Algorithm," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 2459996, 13 pages, 2022.

## Research Article

# Optimal Defect Detection and Sensing System of Railway Tunnel Radar considering Multisensor System Combined with Active Interference Suppression Algorithm

Yang Lei <sup>1</sup>, Yong Zou <sup>2</sup>, Bo Jiang,<sup>1</sup> and Tian Tian<sup>1</sup>

<sup>1</sup>China Academy of Railway Sciences, Beijing 100081, China

<sup>2</sup>Beijing ZhuoRuiXinDa Technology Co., Ltd, Beijing 101599, China

Correspondence should be addressed to Yong Zou; [zouyong@live.cn](mailto:zouyong@live.cn)

Received 31 January 2022; Revised 22 February 2022; Accepted 17 March 2022; Published 25 April 2022

Academic Editor: Daqing Gong

Copyright © 2022 Yang Lei et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the rapid development of science and technology, testing equipment and testing methods are constantly updated. Radar detectors have the advantages of losslessness, high efficiency, high resolution, and high-speed radar image capture. They can accurately locate defects in railway tunnels, respond to hidden dangers in time, and provide strong technical support for transportation. This paper proposes to optimize the defect detection of railway tunnel radar through the combination of multisensor technology and active interference suppression algorithm and designs the corresponding sensor system according to the content. This article analyzes several factors that affect the radar detection effect and makes a detailed summary from the detection environment and other aspects. At the same time, it uses the multisensor system combined with active interference suppression algorithm to design a railway tunnel detection simulation experiment. Experimental results show that the use of multisensors combined with active interference suppression algorithm to optimize radar detection can effectively improve the accuracy of railway tunnel defect detection. Through the analysis of the results of tunnel defect detection, the detection accuracy of this paper has reached 98.8%, which can provide an effective reference for the detection of railway tunnels.

## 1. Introduction

The current railway tunnel construction is a leapfrog development. There are many railway projects in the country under construction. These railways will be the most efficient bridge connecting the country in the future, and the economy will grow rapidly. Therefore, if the construction period is long, on-site construction will often ignore the physical quality of the project. Examples of problems in the operation of railway tunnels are not uncommon in news reports. The lessons are profound, and the results are tragic. Of course, there are many reasons for this tragedy. Everyone involved in the construction is responsible, but in the end, it is due to quality problems. Most of the accidents are caused by the phenomenon of cutting corners and reducing construction procedures during the construction process by the specific construction team at the construction site. At the

same time, the competent authority does not supervise such issues nor do a good job of source management that directly affects the quality of the project. It is precisely because many lessons have been learned that the government attaches great importance to the safe operation of railways. This of course puts forward higher requirements for construction quality. The ground penetrating radar method can effectively solve these problems and can guarantee the safety of the railway in operation for future operations.

There are many methods for tunnel quality inspection. The traditional method is to obtain target information by sampling through boreholes. The detection results obtained by the excavation sampling method are accurate and direct, but only partial information represented by the tunnel lining excavation site can be obtained, and the structure itself has been damaged. Therefore, the quality control of tunnel construction requires an efficient and economical method. It

provides continuous detection information without damaging the tunnel lining itself. Since the 1980s, ground penetrating radar detection technology has been expanding its application range due to its fast, nondestructive, and continuous characteristics. However, the complex environmental conditions of the tunnel under construction and many interference sources will affect the effectiveness of radar wave detection at the detection point. Some electromagnetic interference waves generated by these sources directly affect the reflected signal and the waveform of the target body, and it is difficult to remove them in the post-processing process. In the field test, it should try to avoid these interference sources. Or, it should choose appropriate equipment parameters according to the electromagnetic environment of the construction site to minimize the impact. At the same time, there are other interference factors that affect the accuracy of the detection results, such as environmental humidity, antenna speed, and coupling conditions. The use of ground penetrating radar to comprehensively inspect the tunnel lining structure, finding quality defects and hidden safety hazards and dealing with them in a scientific and timely manner, can avoid serious problems that may even affect the overall safety of the tunnel. It lays a solid foundation for the later completion acceptance and safe operation. It can provide good inspection results and reliable baseline data for project quality evaluation and processing.

In order to overcome active interference, Zhang and Pan proposed an adaptive polarization cancellation method based on dual-polarization radar. They obtain the components and characteristics of the actual target and active interference analysis signal through the vertical and horizontal polarization channels of the dual-polarization radar. Then, in an actual scenario, the weighting factor is subtracted in time to adjust the amplitude and phase difference of the active interference signals received from the two orthogonally polarized channels. This can completely eliminate active interference, but the actual target power will change. In addition, a target signal correction method has been developed to improve the accuracy of monopulse radar angle measurement. Then, through simulation, the factors affecting adaptive polarization cancellation are analyzed in detail. Finally, the experimental results show that it is very effective for the intervention of active blocking and active deception. However, the adaptive method they designed is not flexible enough to meet the actual detection requirements [1]. Zhao et al. researched electronic countermeasures (ECMs) systems to combat misleading electronic countermeasures (ECMs) technology. An adaptive detector uses the generalized probability ratio test (GRLT) standard to detect whether there is any misleading interference in the Fourier transform (FT) field. The proposed detector can analyze initial and auxiliary data through adaptive area echo and misleading interference. Finally, the extended radar network is analyzed as a special case, and the result proves the validity of the digital simulation in the proposed method. However, the scheme they studied is only for special interference [2]. Liu et al. studied the influence of DRFM active jamming technology on coherent pulse compression radar. By

introducing the radar signal detection model and radar equation, the influence of the signal generated by DRFM on the detection performance of the radar receiver is analyzed and the relationship between the radar detection probability and the interference power is obtained. Finally, the effective indicators for judging the interference phenomenon are summarized, and the amplitude of the active interference signal is determined. However, their research on the accuracy of radar detection is not deep enough [3]. Yang et al. proposed an error control method based on an active interference suppression control algorithm. It is used to check the parameter uncertainty and noise measurement of the ultrasonic flutter suppression model. The control method can be used to estimate the observed output signal and its time derivative, as well as to model dynamics and measurement noise. Numerical results show that the control method they invented can effectively suppress ultrasonic vibration in the range of large Mach numbers and has obvious adaptability to changes in stiffness parameters. However, the actual suppression effect of the interference suppression algorithm they studied cannot satisfy the research on geology [4]. Chen et al. proposed a new anti-jamming algorithm based on the minimum mean square error criterion to protect the acquisition stage. The basic idea of the proposed method is the grid search process of sharing synchronization parameters by suppressing interference and receiving signals. The simulation results show that the proposed method can keep the GNSS receiver running during the cold start period with interference signals. In addition, its performance is significantly better than the power reversal method. However, this research method has little meaning for radar detection effect [5]. Yang and Huang studied the interference suppression phenomenon of the integrated coil inductor topology. In the proposed topology, a traditional single-turn spiral inductor is integrated with a built-in tightly coupled LC resonator on the printed circuit board. Interference suppression is achieved by guiding the magnetic flux emitted by the spiral inductor to the tuner instead of other spiral inductors around it. From the experimental results of the comparative case, by changing the value of the coordinator capacitor, the proposed topology can provide design freedom in any frequency band of interest. However, the interference suppression method they proposed is very difficult in practical applications and needs to be simplified [6].

The innovation of this paper is the use of multiple sensors combined with active interference suppression algorithms to optimize the design of a sensor system for radar defect detection in railway tunnels. It uses a multisensor combination scheme to obtain information such as the angular velocity, acceleration, and magnetic field strength of the carrier. A specific calibration compensation method is selected to calibrate their output values, respectively, which realizes the real-time display of system output data. In this article, an interference suppression algorithm based on Fourier transform is introduced to optimize the interference elements encountered by the radar in the detection of radar. The algorithm first transforms the received data in the Fourier transform fractional domain and then separates the

interferometer in the transform domain. By optimizing the design of the gain coefficient, the retention of interference limits is suppressed while signaling time and the frequency response is maximized. Finally, the useful signal after interference suppression is obtained through a short-time inverse Fourier fractional transform. It effectively verifies the effectiveness and simplicity of this method.

## 2. Multisensor System Combined with Active Interference Suppression Algorithm

*2.1. The Working Principle of Radar.* The essence of radar is an electromagnetic sensor. The basic principle is to send an electromagnetic signal to the area of interest through an antenna [7–9] and judge whether there is a target at the end. At the same time, it receives interesting information such as the distance, radial velocity, angular direction, shape, and bandwidth of the target [10, 11]. The radar system can use this information to perform target detection, ranging, tracking, and imaging functions, as well as the process of determining whether the target exists, that is, radar target detection [12, 13].

Ground penetrating radar (GPR) is generally a method of electromagnetic technology [14, 15]. Compared with other detection methods, the geological radar method can directly see the reflected wave recorded by the radar and can directly analyze the distribution and shape of the structure and target in the tunnel [16, 17]. Combining with the preliminary judgment of the target nature and the detection requirements such as detection depth and minimum resolution, it selects antenna types with different frequencies for the highly selective antenna [18, 19]. The ground penetrating radar will not perform destructive testing on the lining surface, so the same section can be tested repeatedly without damaging the tunnel lining structure. The flexible ground penetrating radar has simple operation, high sampling rate, and fast detection speed. Once the speed and analysis ability reach the centimeter level, it can meet most of the requirements of tunnel structure and target detection, with excellent accuracy [20–22]. Its working principle is shown in Figure 1.

*2.2. Active Interference Suppression.* Active interference suppression generates realistic false signals in the radar's time domain, frequency domain, or image field by emitting interference signals that are very close to the echo signal, which affects the radar's detection of actual targets [23]. For active interference suppression, there are countermeasures from the radar system level and signal processing level [24]. However, most of these methods have active suppression. In order to effectively subdue the active suppression of the radar, it is necessary to study how to detect the active radar to suppress the interference [25].

The biggest difference between suppressive interference and deceptive interference is the degree of interference. It is difficult for the radar to obtain effective echo information because it can fully capture the true time and frequency and convert the echo. If the interceptor is powerful enough, it can completely paralyze the radar, making it impossible to fight against [26]. Active interference suppression mechanisms can

be divided into noise amplitude interference, radio frequency noise interference, noise frequency interference, and smart noise interference. According to the implementation strategy of actively suppressing interference, it can be divided into remote support intervention and self-defense intervention [27]. Figure 2 shows the classification of radar interference.

In the case of passive interference, the radar cannot work normally, mainly because the electromagnetic waves emitted by the radar are reflected or changed through the smoke screen. This kind of interference has a good interference effect, a wide range of influence, low cost, and a wide range of application scenarios. Compared with other types of suppressive interference, the interference effect of noise amplitude interference is smaller and its countermeasures are relatively complete [28]. Noise frequency interference is a new type of interference that mainly affects the chip signal. It can automatically target the center frequency of the radar and has a certain pulse compression gain advantage. In recent years, it has become the focus of interferometer technology research. This study mainly analyzes the differences between different types of interference through the modeling and extraction of feature suppression interference and lays the foundation for subsequent suppression of interference detection and identification. Figure 3 shows the process of obtaining target information from the radar.

*2.2.1. Active Interference Suppression Characteristics.* Among the interference factors that affect the radar detection effect, the interference wave generated by the electromagnetic interference object under the action of the electromagnetic field will have a corresponding impact on the radar imaging map. It will cause the imaging spectrum to have no fixed law and range of action. For active interference suppression, extracting its interference characteristics can lay the foundation for radar active interference suppression identification. The following analyzes the six interference patterns to understand their characteristics in the time domain and frequency domain. They extract several characteristic parameters of interference.

(1) *Aiming Frequency Interference.* The time domain expression of noise FM interference is

$$R(a) = Q_i \exp \left[ i \left( 2\phi f_i a + 2\phi H_{f_a} \int a(\varepsilon) b\varepsilon + \kappa_0 \right) \right]. \quad (1)$$

Here,  $Q_i$  indicates that the magnitude of the interference is a constant,  $f_i$  is the carrier frequency that suppresses the interference, and  $H_{f_a}$  is the frequency modulation slope of the FM noise.

The power spectral density of the frequency modulation function  $a(\varepsilon)$  can be expressed as follows:

$$H(\kappa) = \begin{cases} \frac{\phi_n^2}{\Delta\kappa}, & 0 < \kappa < \Delta\kappa, \\ 0, & \text{other,} \end{cases} \quad (2)$$

where  $\Delta\kappa$  is the modulation bandwidth of the modulation noise  $a(\varepsilon)$ .



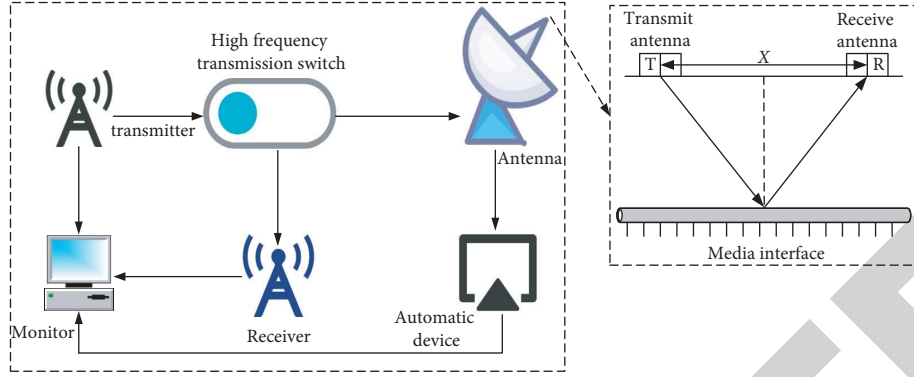


FIGURE 1: Radar working principle diagram.

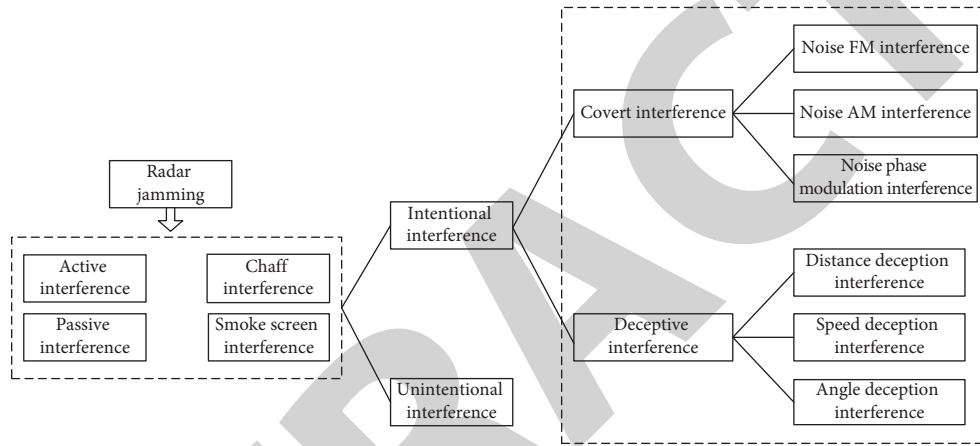


FIGURE 2: Radar jamming classification.

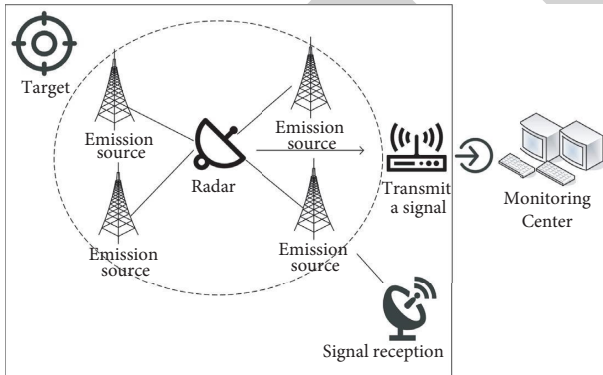


FIGURE 3: The process of radar acquiring target information.

The effective modulation index of the noise FM signal is defined as follows:

$$n_{ke} = \frac{H_{fa}\phi_n}{\Delta\kappa}. \quad (3)$$

Aiming frequency interference is a kind of noise frequency modulation interference, and its center frequency, interference bandwidth, and effective modulation index generally meet the following conditions:

$$\begin{aligned} f_i &= (\phi f + f_0), \\ \Delta\kappa_i &= (2 \sim 5)\kappa_r, \\ n_{fa} &< 0. \end{aligned} \quad (4)$$

Here,  $f_0$  represents the center frequency of the radar signal,  $\phi f$  represents the aiming frequency error,  $\Delta\kappa_i$  represents the bandwidth of the aiming interference signal, and  $\kappa_r$  represents the bandwidth of the radar signal. Target frequency interference usually requires that the frequency of the interferometer matches the center frequency of the radar. At the same time, it transmits interference signals with a narrow interference bandwidth, which covers the receiving bandwidth of the radar and requires a high-precision frequency detector to support interference. However, in recent years, with the development of radar frequency flexibility technology, it has become difficult to identify the center frequency of the radar from the interference part and it may be difficult to intercept the target frequency interference on the radar by using frequency interference.

(2) *Blocking Interference*. Blocking interference is a type of FM noise interference. The interference bandwidth and effective configuration index must usually meet the following conditions:

$$\begin{aligned} f_i &= (\phi f + f_0), \\ \Delta\kappa_i &= 5\kappa_r, \\ n_{fa} &> 0. \end{aligned} \quad (5)$$

(3) *Sweep Frequency Interference*. Sweep frequency interference can be regarded as a kind of aiming frequency interference whose center frequency is constantly changing. Under normal circumstances, this kind of interference needs to meet the following conditions:

$$\kappa_i = \phi f + f_0 + \frac{\Delta\kappa_b}{a} \times A - \frac{\Delta\kappa_b}{2} a \in [0, a_b]. \quad (6)$$

Here,  $\Delta\kappa_b$  is the bandwidth of the frequency sweep interference and  $a_b$  is the frequency sweep period of the frequency sweep interference.

The bandwidth of frequency sweep interference is very large, and it can compete well with flexible-frequency radars and multiple radars of different frequencies. If the given frequency bands are the same, the aggressive interference power will be higher than the cancellation interference power. However, this interference becomes discontinuous in the tie domain after passing through the receiver. The interference scanning frequency represents the number of times a scanning interference scans the entire interference bandwidth per unit time, expressed as follows:

$$\Delta\kappa_i = JSF = \frac{1}{A_b}. \quad (7)$$

Sweep frequency interference will form intermittent pulse interference after passing through the receiver, and the larger the pulse, the denser the pulse.

(4) *Comb Spectrum Interference*. Comb spectrum interference is an interference pattern composed of multiple narrowband interferences, and its expression can be summarized as follows:

$$R(a) = \sum_{i=1}^m R_i(a) + \frac{\Delta\kappa_a}{A} \quad i = 1, 2, \dots, m, \quad (8)$$

where  $m$  represents the number of narrowband interferences included in the spectrum interference. Each narrowband interference is independent of each other. There are different frequency points in the interference center, and spectrum interference can produce narrowband interference at multiple different frequency points. It can successfully deal with variable frequency radar systems such as flexible-frequency radars and frequency-fluctuation radars.

(5) *Noise Product Interference*. Noise product interference is a new type of interference directed at chirp signals. Its mechanism is to multiply the noise through the filter and the radar pulse signal intercepted by the jammer. This interference can be expressed as follows:

$$R(a) = m(a) \times b(a) \quad i = 1, 2, \dots, m, \quad (9)$$

where  $m(a)$  is the noise passing through the filter and  $b(a)$  is the chirp signal intercepted by the jammer, and its expression is

$$R(a) = \text{rect}\left(\frac{1}{A}\right) \exp\left(i2\phi\left(f_0 + \frac{1}{2} \times \frac{G}{A} a^2\right)\right). \quad (10)$$

The variable  $f_0$  represents the center frequency of the chirp signal,  $G$  represents the bandwidth of the FM signal, and  $A$  represents the time width of the FM signal. The spectrum of noise product interference can be expressed as follows:

$$I(f) = M(f) \otimes B(f). \quad (11)$$

(6) *Noise Convolution Interference*. Noise convolution interference is also a new type of interference for chirp signals. Its mechanism is to convolve the noise through the filter with the radar pulse signal intercepted by the jammer. This interference can be expressed as follows:

$$R(a) = m(a) \otimes B(f) a > 0. \quad (12)$$

It can be seen from the representation of the convergence noise interference frequency domain that the convergence noise interference coincides with the center frequency of the transmitting radar signal and the bandwidth is close; that is, the convergence noise interference can be aimed at the center frequency radar. Without precise frequency measurement, it can effectively affect flexible-frequency radars, frequency-differential radars, and other radar systems that allow flexible radar signal ranges. The essence of convergent noise interference can be thought of as amplifying the received chirp signal at different rates and then delaying the addition. Therefore, convergence noise interference can achieve a specific radar pulse compression gain, that is, after the interference is processed by the noise convolutional radar system, it can achieve the same interference effect as other low-power interference suppression.

**2.3. Interference Suppression Algorithm.** The analysis of active radar interference suppression mainly starts from the feature derivation. Various functional parameters are derived and analyzed from the time domain, frequency domain, and transform domain to lay the foundation for subsequent active radar interference detection. The interference characteristic parameters are as follows.

**2.3.1. Frequency-Domain Peak-to-Average Power Ratio after the Receiver.** Assuming that the radar echo signal is  $v(a)$ , the receiver noise is  $d(a)$  and the received interference is  $i(a)$ , then the radar received signal  $w(a)$  can be expressed as follows:

$$w(a) = v(a) + d(a) + i(a). \quad (13)$$

After sampling  $w(a)$ , perform the fast Fourier transform to obtain the following:

$$w_i(n) = \frac{w(n)}{\max(w(n))_{n=1}^n} \quad n = 1, 2, \dots, N. \quad (14)$$

The frequency-domain peak-to-average power ratio after the receiver is defined as follows:

$$P_i(a) = 10 \log_{10} \frac{\max_{n=1}^N w_i^2(n)}{1/N w_i^2(n)}. \quad (15)$$

The peak-to-average power ratio parameter in the frequency domain mainly reflects the signal variation range in the frequency domain. Its value is relatively high, and the range of comb spectrum interference can be identified by this feature. The flow of the interference suppression algorithm is shown in Figure 4.

The echo signal is divided into several time periods at equal intervals in the time domain. First, FT is applied to the first time period data, and the optimal order is obtained by searching. By analogy, the optimal order of each time window is obtained. After Fourier transform, the time-frequency aggregation of narrowband and broadband interference is obtained. Then, this is used to detect each time slice. If there is an interference signal, an adaptive gain control method is used to suppress the interference signal to complete the interference suppression.

We calculate the initial value of the argument of  $w(n)$ , adopt the principle of the main value interval, and use the optional interval  $(-\sigma \sigma)$  as the main value interval to ensure the odd symmetry of the argument with respect to  $n=0$ . To eliminate the phase ambiguity of logarithmic changes, there are

$$\ln(w(a)) = \ln(I(a)) + \ln(1 + d'(a)). \quad (16)$$

It can be seen from the formula that the real part and the imaginary part are, respectively,

$$\begin{aligned} R_w(a) &= ie[(w(a))] = \ln W_i + ie[d'(a)], \\ I_w(a) &= \ln[\ln w(a)] = \sigma_i + \ln[d'(a)]. \end{aligned} \quad (17)$$

Since both the deception interference and the echo signal are zero-average signals, the amplitude of the noise FM interference signal can be estimated as follows:

$$\widehat{W}_i = \exp\{\text{mean}[R_w(a)]\}. \quad (18)$$

We calculate the interference signal  $i(a)$  from the radar received signal  $w(a)$ ; then,

$$\begin{aligned} \widehat{d}(a) &= w(a) - \widehat{I}(a) \\ &= d'(a) + I(a) \left[ 1 - \frac{\widehat{W}_i}{W_i} \exp(i \ln(d'(a))) \right]. \end{aligned} \quad (19)$$

### 3. Experiment of Railway Tunnel Radar with the Multisensor System Combined with Active Interference Suppression Algorithm

As a kind of concealed project, tunnel construction will inevitably have defects in the construction process. Due to its concealment, defect detection has become a concern of tunnel workers. Geological radar detection is a fast and efficient detection method. Its principle is the same as that of the tunnel lining structure. The tunnel lining structure is combined with radar detection for tracking tunnel defects.

The ground penetrating radar is used for nondestructive testing of tunnel lining, and its main purpose is to ensure that the construction quality of tunnel lining meets the requirements of design specifications. The tested elements are the initial support, secondary lining and inverted arch, concrete thickness detector, double-layer secondary lining thickness, gaps and nonsolid areas, and concrete compression of the backfill layer.

Due to the unevenness of the coated concrete, when the electromagnetic wave propagates to the initial support and the secondary coating, the speed of the electromagnetic wave will change within a certain range. We need to adjust the parameters of the medium to be measured on the concrete. If we want to calibrate the relative permeability or electromagnetic velocity of shotcrete before the first support material inspection, we can use the core perforation sampling method to calibrate three or more positions and take the average value. In this paper, the tunnel is inspected in the part where the concrete thickness is known, because the core drilling sampling method will damage the back wall to calibrate the parameters of the secondary cladding concrete. The parameters of the radar include the number of sampling points, sampling rate, and band-pass filtering. The number of sampling points represents the number of sampling points in each acquisition dataset. This article sets the number of sampling points to 400. In order to make the recorded waveform more complete and controllable, the sampling rate is selected to be 6 times the antenna center frequency. In the processing of data and interference waves, band-pass filtering is used.

Due to the requirement of a safe step between the surface and cladding of the tunnel under construction, the construction of the tunnel cladding should be monitored quickly. Therefore, the initial support inspection should monitor the construction progress. It should take into account the inspection progress, that is, whether the age of the concrete can meet the inspection requirements. Before the inspection, the side walls of the cladding should be clearly marked with red paint every 5 meters, indicating the number of kilometers and the number of piles. If necessary, the detection location should be clearly marked to ensure the accuracy of the detection data.

The main control measures for the secondary cladding inspection of the tunnel are used for cavities, defects, cladding thickness, inverted thickness, filler layer and other cladding, the distribution of steel bars in the cladding, and the solid lining of the initial support. The main detection method is the geological radar method. The tunnel wall inspection is to place inspection lines on the tunnel warehouse, the center arch on both sides, the side walls, and the inverted arch on both sides. Or, it provides a basic overview of the cross section of the intersection. And, it focuses on the encrypted overview of specific tunnel sections or irregularly covered structural sections, as well as gives the basic knowledge of paving leaks, horizon cracks, and crevices. Figure 5 is a schematic diagram of radar defect detection in railway tunnels.

The tunnel defect detection system requires methane sensors, temperature and humidity sensors, wind speed and



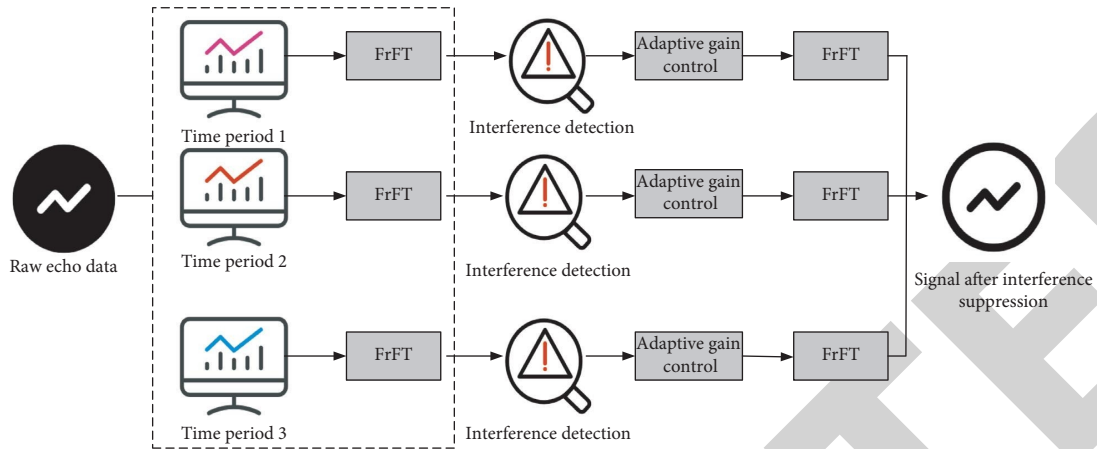


FIGURE 4: Flow chart of interference suppression algorithm.

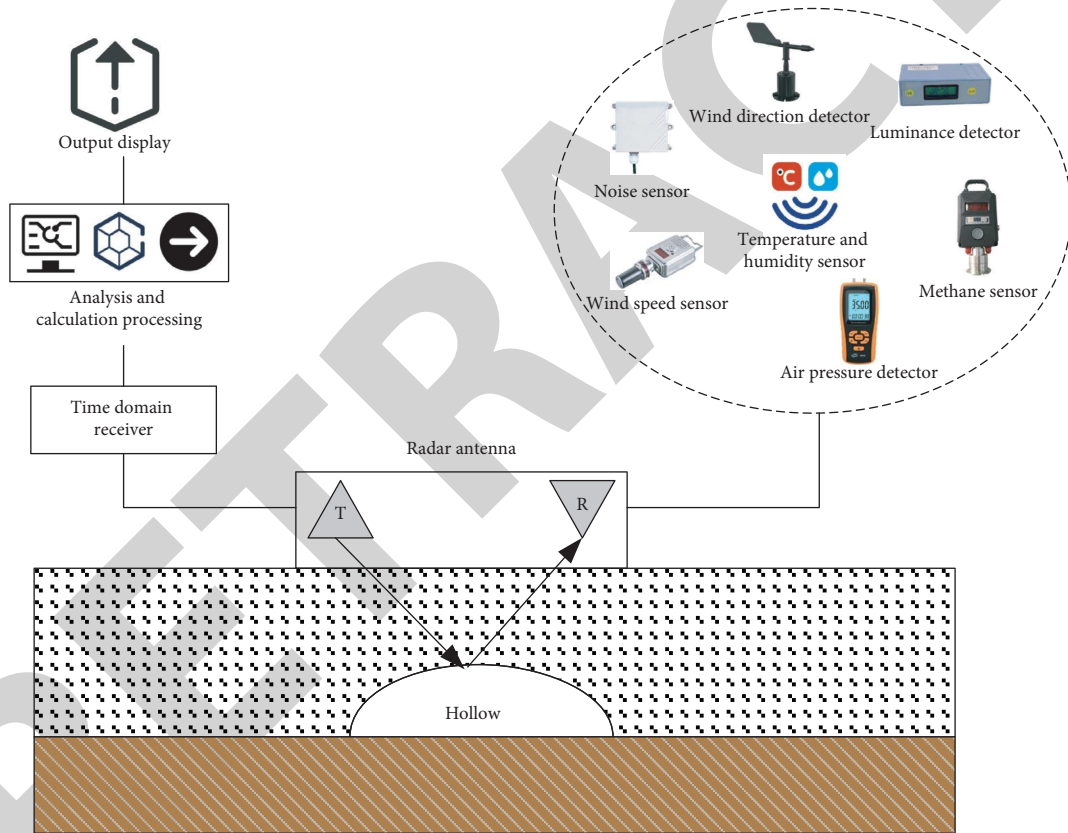


FIGURE 5: Schematic diagram of railway tunnel radar defect detection.

direction sensors, noise sensors, brightness detectors, and air pressure detectors. At the same time, it cooperates with radar to detect railway tunnels. And, it provides statistics on the size and location of defects behind the tunnel lining and summarizes the distribution of specific types of defects. The defect classification table is now introduced. Table 1 is the tunnel defect classification table.

In the detection of railway tunnel defects, firstly, the relevant sensors in the defect detection system are used to obtain the data related to the tunnel defects at the moment and then the output data of the sensors are processed

accordingly. In this process, any calculation processing is based on sensor data. Therefore, the sensor plays a vital role in the tunnel defect detection system. Different sensors have different test principles and applicable conditions. These will affect the results of tunnel defect detection and the application environment of the system. Figure 6 is the design of a multisensor system for railway tunnel defect detection.

The noise sensor, temperature and humidity sensor, wind speed and direction sensor, and methane sensor are combined with a high-precision microprocessor as the control and calculation unit to obtain the relevant

TABLE 1: Tunnel defect classification table.

Defective item	Defect level	I	II	III	IV
		Slight defect length (m)	More serious defect length (m)	Severe defect length (m)	Very serious defect length (m)
There is a cavity behind the lining	There is a hole	$\leq 1$	$1 \leq L \leq 3$	$3 \leq L \leq 5$	$> 5$
Backfill is not dense	Not dense	$\leq 3$	$3 \leq L \leq 9$	$9 \leq L \leq 15$	$> 15$
The base is not dense	Not dense	$\leq 3$	$3 \leq L \leq 9$	$9 \leq L \leq 15$	$> 15$

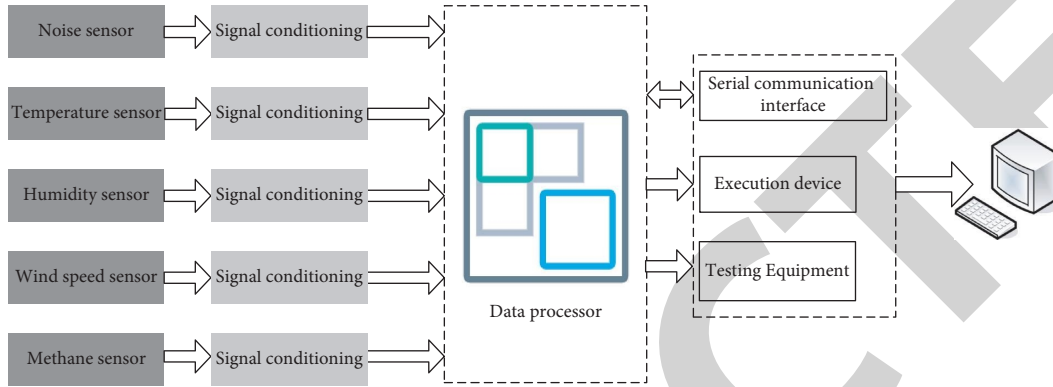


FIGURE 6: Design of the multisensor system for railway tunnel detection.

information of the railway tunnel and obtain the environmental factors of the tunnel. Then, it cooperates with the air pressure detector, light brightness detector, and magnetic field detector to obtain the air pressure, light brightness, and magnetic field strength in the tunnel.

#### 4. Experimental Results and Analysis

**4.1. Simulation Analysis of the Multisensor System Combined with Active Interference Suppression Algorithm.** In order to analyze the algorithm more deeply, simulation experiments are carried out using the algorithm. Figure 7 is a histogram of phase statistics. It can be seen from the figure that the echo is widely distributed in the entire phase interval, and the echo and interference have obvious different phase distribution characteristics. In contrast, the interference is only concentrated on a limited phase. Therefore, the interference and echo can be identified through the conventional discrimination model of the tree.

Figure 8 shows the output waveforms of the radar-matched filter before and after interference suppression. It can be seen from Figure 8(a) that before the interference suppression of the matched filter, both the target and the interference spike exist in its output. Therefore, the radar cannot achieve the purpose of extracting real target information by selecting target spikes from multiple spikes. In contrast, the matched filter output in Figure 8(b) has only a target spike after interference suppression, and the radar can extract the distance information of the real target through this spike.

Figure 9 is a performance graph of the radar suppression algorithm. Under various conditions, the target echo similarity curve is consistent with the probability curve of correct target recognition after radar interference. It can be seen from this figure that as the signal-to-noise ratio

increases, the interference and echo similarity coefficient increases. And, the phase is sensitive to the influence of noise under low signal-to-noise ratio, and it is not easy to be distinguished.

Figure 10 is the comparison result of several methods under different frequency deviations. It can be seen from the figure that when the frequency estimation performance of the improved method reaches the optimum, the frequency deviation in the constraint condition can be satisfied. At this time, the frequency estimation accuracy of the improved method is significantly higher than that of the other three methods.

**4.2. Defect Detection Problems in Railway Tunnels.** As an important nondestructive testing method, radar has been accepted by engineers. At present, it is widely used in the inspection of heavy and large projects at home and abroad, and it controls the quality of the project very well. However, due to various factors, there are more or less problems in the application of geological radar. In response to these problems, the application scholars of the ground penetrating radar have done a lot of research in order to reduce the detection error of the ground penetrating radar. This article specifically studies and analyzes tunnel lining inspection and lists the factors that affect tunnel inspection, so as to improve the accuracy of domestic tunnel lining inspection. It provides safety guarantee for tunnel engineering.

Concrete is a relatively stable synthetic material. It is a nonelectric good conductor, that is, a nonconductive substance. From Table 2, we can see that the dielectric properties of concrete will also have an impact due to the difference in water content.

As shown in Figure 11, according to the dielectric constant results of concrete at different ages, the dielectric

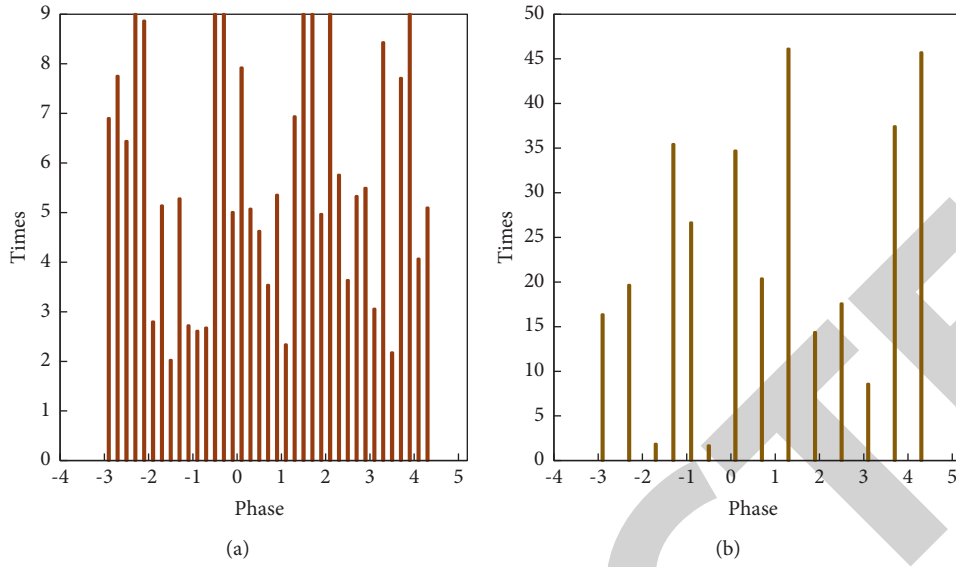


FIGURE 7: Source signal phase statistics histogram.

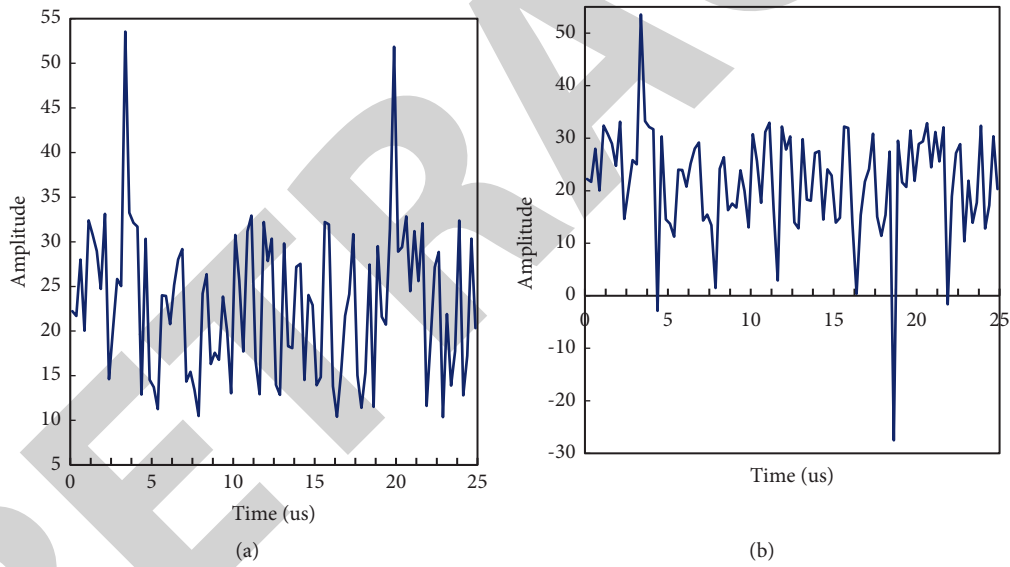


FIGURE 8: Radar-matched filter output waveform (a) before and (b) after interference suppression.

constant of concrete at 3d, 7d, 14d, and 28d age was analyzed.

In addition to the relative permittivity and permeability of the medium, the vertical resolution of the ground penetrating radar is also related to the center frequency of the antenna. When the medium is constant, the vertical resolution of the ground penetrating radar is inversely proportional to the center frequency of the radar antenna. Regarding the influencing factors, radar should pay attention to the selection of detection timing when detecting defects. At the same time, when facing electromagnetic interference, it should be far away from these interference sources or appropriate shielding measures should be selected to minimize the impact.

**4.3. The Detection Results of Railway Tunnel Radar Optimization.** In the process of detecting defects in railway tunnels, electromagnetic waves first pass through the air medium and enter the concrete medium. For air and concrete media, the reflection coefficient is negative. In other words, electromagnetic waves are negative. Therefore, the first wave of the measurement results is inconsistent and the difference is large. Table 3 shows the measurement location information of the first wave.

Figure 12 is the detection accuracy and error analysis results of the railway tunnel radar defect detection of the multisensor system combined with active jamming algorithm.

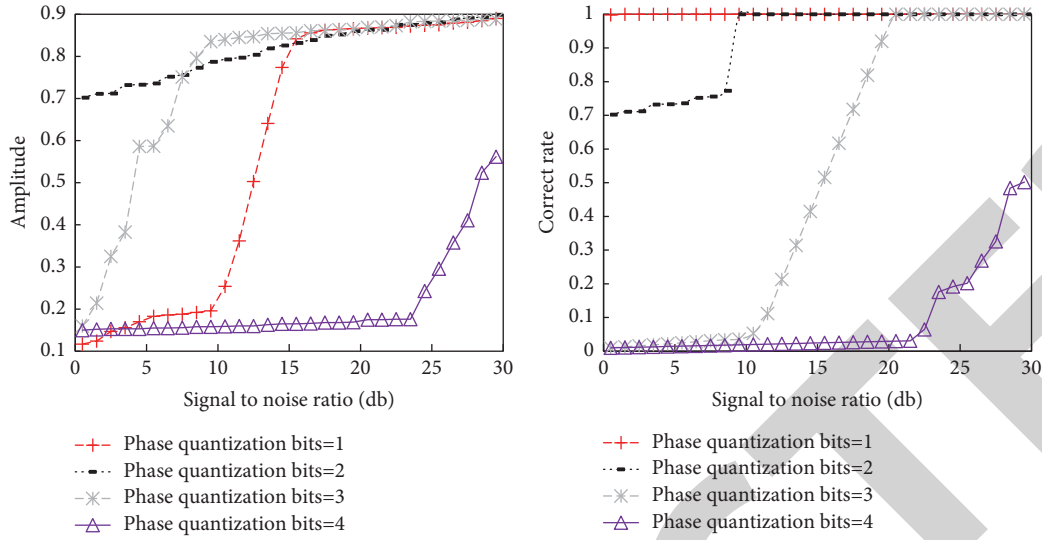


FIGURE 9: Radar suppression algorithm performance graph.

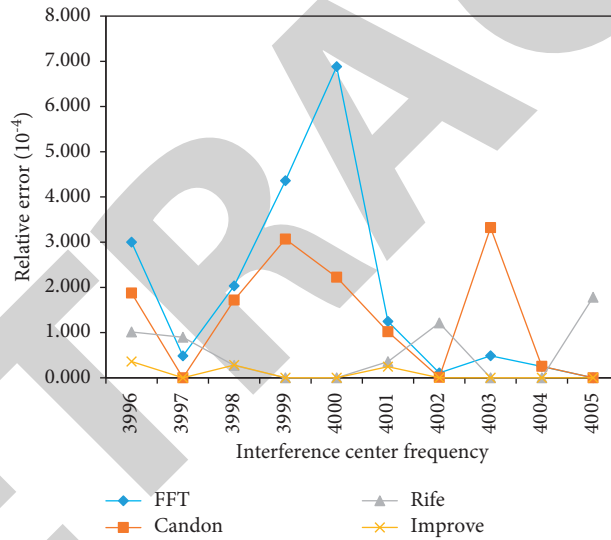


FIGURE 10: Comparison results of several methods under different frequency deviations.

TABLE 2: The difference between the dielectric constant of each parameter and the thickness of the test.

Dielectric constant		Volume ratio%		Concrete dielectric constant	Calculated lining thickness
Air	Water	Air	Water		
1	81	5	1	6.25	0.38
1	81	4	2	6.74	0.37
1	81	3	3	7.32	0.36
1	81	2	4	7.68	0.35
1	81	1	5	8.31	0.34

It can be seen from the figure that the traditional detection method can easily make the cladding thickness too large or too small, resulting in a large error of up to 0.042 m in the identification and evaluation of railway tunnel defects. It greatly reduces the reliability and is easy to make wrong estimates or cause unnecessary losses. The detection method

optimized by combining multiple sensors and active interference suppression algorithm can confirm that the error is within 0.02 m. The multisensor defect detection rate combined with the active interference suppression algorithm is 98.8%, which can effectively improve the detection accuracy.

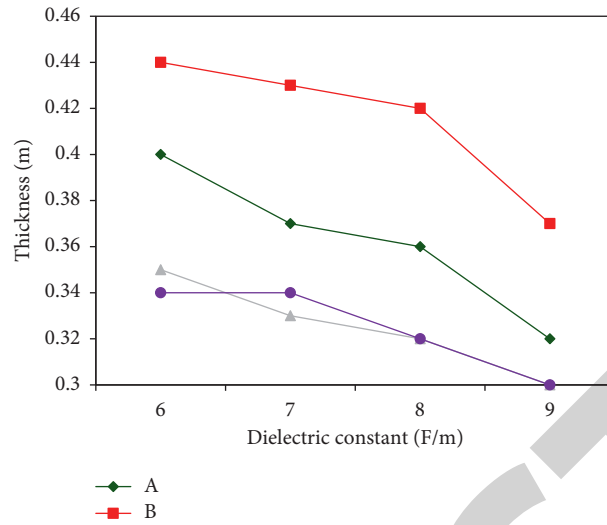


FIGURE 11: The lining thickness of concrete of different ages.

TABLE 3: First wave signal position table.

Number	First wave signal position (ns)	Mean difference (ns)	Maximum difference (ns)	Minimum difference (ns)
1	3.9	-0.23	0.22	-0.18
2	3	0.09	0.28	-0.32
3	3.16	0.17	0.16	-0.11
4	3.18	0.14	0.14	-0.11
5	3.71	0.09	0.13	-0.16
6	3.83	-0.12	0.22	-0.05
7	3.34	-0.17	0.07	-0.24
8	3.83	0.27	0.19	-0.11
9	3.66	0.18	0.11	-0.15
10	3.78	0.21	0.14	-0.12

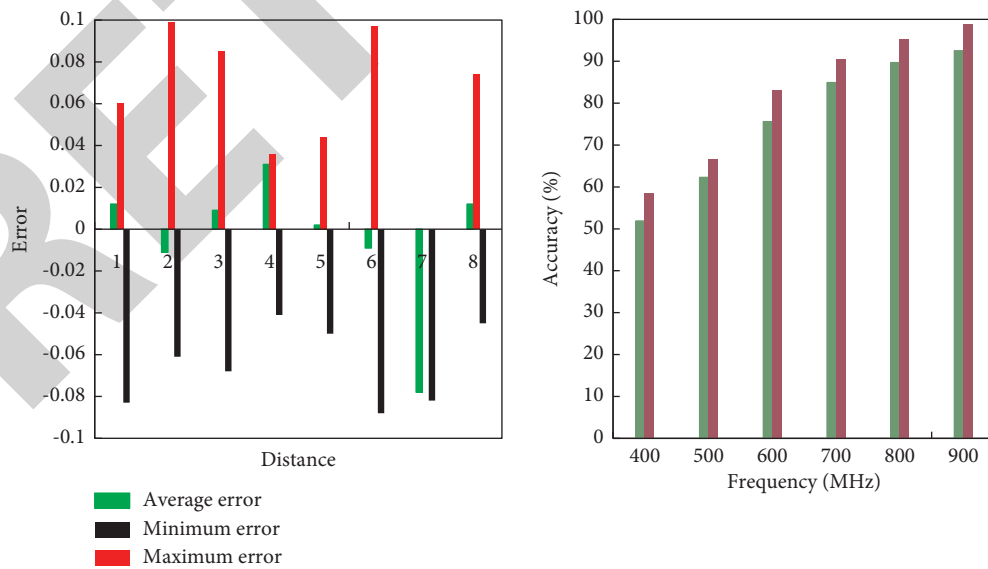


FIGURE 12: Detection accuracy and error.

## 5. Conclusion

In this paper, the defect detection of railway tunnels is optimized with the multisensor system combined with active interference suppression algorithm and a related sensor detection system is designed. The multisensor detection system includes temperature and humidity sensors, wind speed and direction sensors, methane sensors, and noise sensors. It can meet the needs of tunnel curve detection, and the multisensor detection system has higher stability and accuracy. It starts from the basic principles of ground penetrating radar and electromagnetic field theory, according to the propagation characteristics and laws of electromagnetic waves and various tunnel radar detection parameters and indicators, to study data acquisition, processing, and transmission methods. The results of radar detection factors of railway tunnels include the influence of equipment parameters and the influence of the external detection environment. In this paper, combined with relevant examples of mechanical substance detection, the specific application of ground penetrating radar in various factors that affect the quality of tunnel detection, matters needing attention, railway tunnel radar detection, etc., is analyzed. It focuses on analyzing the amplitude and polarity of the reflected wave, the spectral characteristics of the reflected wave, the morphological characteristics of the internal axis of the reflected wave, and the influence of interference waves. At the same time, this paper studies the influence of water and air on the dielectric constant of concrete and the influence of various factors on the radar detection, such as the choice of detection timing when detecting the thickness of concrete. Finally, it is hoped that the content studied in this article can provide useful information and reference value for radar operators.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare no conflicts of interest.

## Acknowledgments

This research was funded by the Key Project of Science and Technology Research of China Academy of Railway Sciences, under grant number 2019YJ157, and Project of Science and Technology Research and Development Plan of China Railway Corporation, under grant number 2017G003-H.

## References

- [1] Q. Zhang and W. Pan, "Countering method for active jamming based on dual-polarization radar seeker," *International Journal of Microwave & Wireless Technologies*, vol. 9, no. 5, pp. 737-738, 2017.
- [2] Y. Zhao, Z. Ran, Y. Xiong, and T. A. Bin, "ABORT-like detector to combat active deceptive jamming in a network of LFM radars," *Chinese Journal of Aeronautics*, vol. 30, no. 4, pp. 1538-1547, 2017.
- [3] W. Liu, J. Meng, and L. Zhou, "Impact analysis of DRFM," *Journal of Engineering*, vol. 2019, no. 20, p. 6858, 2019.
- [4] Z. Yang, R. Huang, Y. Zhao, and H. Hu, "Design of an active disturbance rejection control for transonic flutter suppression," *Journal of Guidance, Control, and Dynamics*, vol. 40, no. 11, pp. 2905-2916, 2017.
- [5] F. Chen, J. Nie, S. Ni, Z. Li, and F. Wang, "Combined algorithm for interference suppression and signal acquisition in GNSS receivers," *Electronics Letters*, vol. 53, no. 4, pp. 274-275, 2017.
- [6] S.-F. Yang and T.-H. Huang, "Design of single-turn spiral inductors with embedding A strong-coupling LC resonator for interference suppression," *IEEE Transactions on Electromagnetic Compatibility*, vol. 59, no. 3, pp. 919-926, 2017.
- [7] L. Zhang, H. Gao, J. Wen, S. Li, and Q. Liu, "A deep learning-based recognition method for degradation monitoring of ball screw with multi-sensor data fusion," *Microelectronics Reliability*, vol. 75, pp. 215-222, 2017.
- [8] W. Xiong, M. Greco, F. Gini, G. Zhang, and Z. Peng, "SFMM design in colocated CS," *IET Radar, Sonar & Navigation*, vol. 12, no. 7, pp. 702-710, 2018.
- [9] A. Lorenc, M. Kužnar, T. Lerher, and M. Szkoda, "Predicting the probability of cargo theft for individual cases in railway transport," *Tehnicki vjesnik-Technical Gazette*, vol. 27, no. 3, pp. 773-780, 2020.
- [10] I. Jang, J. Jeong, H.-S. Shin, S. Kim, A. Tsourdos, and J. Suk, "Cooperative control for a flight array of UAVs and an application in radar jamming \* \*the authors gratefully acknowledge that this research was supported by international joint research programme with chungnam national university (No. EFA3004Z)," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 8011-8018, 2017.
- [11] Z. Xing and Y. Xia, "Distributed federated kalman filter fusion over multi-sensor unreliable networked systems," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 63, no. 10, pp. 1714-1725, 2017.
- [12] S. Subedi, Y. D. Zhang, M. G. Amin, and B. Himed, "Cramer-Rao type bounds for sparsity-aware multi-sensor multi-target tracking," *Signal Processing*, vol. 145, pp. 68-77, 2017.
- [13] Q. Shi, J. Huang, T. Xie, C. Wang, and N. Yuan, "An active jamming method against ISAR based on periodic binary phase modulation," *IEEE Sensors Journal*, vol. 19, no. 18, pp. 7950-7960, 2019.
- [14] D. Adamy, "Jamming semi-active guided threats," *Journal of Electronic Defense*, vol. 42, no. 4, pp. 43-44, 2019.
- [15] D. I. Likhovitskiy, V. P. Riabukha, A. V. Semeniaka, D. V. Atamanskiy, and Y. A. Katiushyn, "Protection of coherent pulse radars against combined interferences. 1. Modifications of STSP systems and their ultimate performance capabilities," *Radioelectronics and Communications Systems*, vol. 62, no. 7, pp. 311-341, 2019.
- [16] R. . Scott, "Saab develops radar mode for hypersonic detection," *Jane's Navy International*, vol. 123, no. 9, p. 8, 2018.
- [17] M. Xie, S. Zhao, X. Fu, T. Zhang, and K. Liu, "Chaff jamming recognition of radar," *Journal of Beijing Institute of Technology*, vol. 95, no. 1, pp. 47-54, 2018.
- [18] W. Yi, M. Jiang, R. Hoseinnezhad, and B. Wang, "Distributed multi-sensor fusion using generalised multi-Bernoulli densities," *IET Radar, Sonar & Navigation*, vol. 11, no. 3, pp. 434-443, 2017.



## Research Article

# Fault Diagnosis in Regenerative Braking System of Hybrid Electric Vehicles by Using Semigroup of Finite-State Deterministic Fully Intuitionistic Fuzzy Automata

Sajida Kousar <sup>1</sup>, Farah Aslam <sup>1</sup>, Nasreen Kausar <sup>2</sup>, Dragan Pamucar <sup>3</sup>,  
and Gezahagne Mulat Addis <sup>4</sup>

<sup>1</sup>Department of Mathematics and Statistics, International Islamic University Islamabad, Islamabad, Pakistan

<sup>2</sup>Department of Mathematics, Faculty of Arts and Sciences, Yildiz Technical University, Esenler 34210, Istanbul, Turkey

<sup>3</sup>Department of Logistics, University of Defence in Belgrade, Belgrade, Serbia

<sup>4</sup>Department of Mathematics, University of Gondar, P.O. Box: 196, Gondar, Ethiopia

Correspondence should be addressed to Gezahagne Mulat Addis; gezahagne412@gmail.com

Received 3 February 2022; Revised 14 March 2022; Accepted 16 March 2022; Published 22 April 2022

Academic Editor: Daqing Gong

Copyright © 2022 Sajida Kousar et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Regenerative braking is one of the most promising and ecologically friendly solutions for improving energy efficiency and vehicle stability in electric and hybrid electric cars. This research describes a data-driven method for detecting and diagnosing issues in hybrid electric vehicle regenerative braking systems. Early fault identification can help enhance system performance and health. This study is centered on the construction of an inference system for fault diagnosis in a generalized fuzzy environment. For such an inference system, finite-state deterministic fully intuitionistic fuzzy automata (FDFIFA) are established. Semigroup of FDFIFA and its algebraic properties including substructures and structure-preserving maps are investigated. The inference system uses FDFIFA semigroups as variables, and FDFIFA semigroup homomorphisms are employed to illustrate the relationship between variables. The newly established model is then applied to diagnose the possible fault and their nature in the regenerative braking systems of hybrid electric vehicles by modeling the performance of superchargers and air coolers. The method may be used to evaluate faults in a wide range of systems, including autos and aerospace systems.

## 1. Introduction

Complex processes and phenomena are prevalent in modern science and technology, about which comprehensive knowledge is not always accessible. Mathematical models are created to address many types of systems having aspects of uncertainty in such instances. A substantial portion of these approaches is built on the so-called fuzzy sets, which are a new extension of conventional set theory. The fuzzy set proposed by Zadeh [1] is based on the formulation of membership function from  $X$  to  $[0, 1]$ , where the images are termed as membership grades or degrees of membership of elements of  $X$ . Atanassov [2] proposed the notion of an intuitionistic fuzzy set (IFS), which is an extension of the perception of the fuzzy set where the degree of

nonmembership is also considered along with the degree of membership. A finite-state machine (FSM) or finite-state automata (FSA) are a mathematical model of computation that can be changed from one state to another state in connection to suitable inputs. There are two types of FSMs, deterministic FSM (DFSM), the one who accepts or rejects a given string of inputs, following a state sequence uniquely obtained from the string, and nondeterministic FSM (NDFSM), which does not obey these restrictions. Moreover, for each NDFSM an equivalent DFSM can be constructed. Fuzzy automata are used to handle system uncertainties more accurately, whereas classical automata fail to cater to the circumstances. Fuzzy automata have been frequently employed since the introduction of fuzzy technology and neural networks [3–13]. Furthermore, there were



a variety of problems to be resolved, for example, medical diagnosis, car anti-crash radar, freeway management, urban road traffic control, and obstacle recognition in front of a vehicle, which required flexible, quick, and accurate decisions, and then, fuzzy neural network automata (FNNA) [14–17] are an excellent choice. FNNA had an increasingly prominent role, particularly in data communications. In FNNA, fuzzy technology is used to compare ordinary identification and control devices using several features and techniques of the neural network including fast and accurate numerical calculation of large numbers and self-organization learning ability. As a result, FNNA was able to solve a variety of problems in a flexible, quick, and precise manner and it had several advantages such as compact size, lightweight, and stability. The research findings also revealed that it has a significant ability for self-organization learning and the experimental results demonstrated fuzzy automata's superiority. Li and Qiu [18] introduced the technique of minimizing fuzzy automata and constructed a new automaton system that dealt accurately with many states at a time. They used an ordered lattice to reduce a fuzzy automaton of  $n$  states to another fuzzy automaton with  $m$  states ( $m < n$ ) that were functionally correspondent to the original fuzzy automata. Several language-preserving methods for minimizing deterministic fuzzy automata were established. Moghari and Zahedi's [19] method preserves both language and structure in minimization process. De Mendivil and Garitagoitia [20] described the factorization of fuzzy states that are used in the determination of fuzzy automata. Stamenkovic et al. [21] studied fuzzy automata and reduction in fuzzy automaton states and relational equations of the fuzzy system by considering the solution of the fuzzy equivalent equation. These solutions are then used to reduce the number of acceptable states. The solution of the fuzzy system of quasi-order number was also used to reduce the states of the fuzzy automaton.

The relationship between semigroup and FSA was established by Krohn and Rhodes' [22] discussed semigroup decomposition employing decomposing of FSA. A semigroup is an important algebraic structure, which serves as the theoretical foundation for a variety of scientific fields with several applications [23]. Semigroup's role in theoretical computer science is inevitable; in particular, semigroup and automata are widely studied and applied in artificial intelligence, game theory, dynamical system, system biology, and fault diagnosis. During the investigation, it was discovered that to diagnose a machine's problem situation, it is necessary to look at its operating state, fault degree, accuracy degree, and expected changes between states at different stages, which can be done by establishing an appropriate FSA. The uncertainty and ambiguity in the diagnostic process enforce to carry out the mathematical computation in a fuzzy environment. A semigroup of finite deterministic intuitionistic fuzzy automata (FDIFA) is used to detect default in any machine [24] by identifying its crucial components, and it can perform characteristic processing using its state transition mapping. Furthermore, because it accurately and efficiently detects the fault, the maximum value of the membership and nonmembership

grades can be used to diagnose the maximum default and current running state of any machine. Membership and nonmembership grades can be assigned based on the nature and seriousness of the fault. The algebraic properties of FDIFA semigroup are employed in the formulation of inference systems over FDIFA semigroup. Fuzzy inference is a method of formulating a rule or a mapping from a given set of inputs to a given set of outputs utilizing fuzzy logic. The fuzzy inference system is widely used in control systems [25], artificial intelligence [26, 27], decision-making [28], forecasting [29], and game theory [30]. Fuzzy set over a non-empty set  $X$  as defined by Zadeh is a function (membership function) from  $X$  to the closed interval  $[0, 1]$  so the representation of uncertain data using closed interval is equivalent to representation using fuzzy sets [31, 32]. The fuzzy inference of fuzzy automata is introduced and utilized in machine defect diagnosis as the impact of combination of fuzzy automaton rules and method of fuzzy set is greater than the classical automata. For the derivation of automata, some researchers used neural networks [33, 34]. Although neural networks and fuzzy systems are not the same, fuzzy inference has been widely applied in the disciplines of control and intelligence [35]. The fuzzy inference system and fuzzy automata and algebraic structures on fuzzy automata are investigated and successfully applied in computer science and engineering [36]. A intuitionistic fuzzy set is a very useful generalization of fuzzy set, which provides an effective framework to handle imprecision and uncertainties more accurately than the fuzzy set and logic. This motivates us to incorporate the concept of nonmembership grades in fuzzy automata. The work is based on the development of inference systems and automata in an intuitionistic fuzzy environment. Intuitionistic fuzzy automata have several states at a certain time or stage, each equipped with a certain grade of reliability and non-reliability. Thus, the intuitionistic fuzzy automata are more realistic and close to the objective. To achieve the goal, finite-state deterministic fully intuitionistic fuzzy automata (FDFIFA) are defined in Section 2. In Section 3, monoid on fully intuitionistic fuzzy automata is designed and its algebraic properties are investigated, and intuitionistic fuzzy inference rules on FDFIFA monoids are established in Section 4. In Section 5, the mathematical computations are supported and justified by conduction fault diagnosis in hybrid vehicle regenerative braking system in application (see Figure 1).

## 2. Fully Deterministic Finite-State Intuitionistic Fuzzy Automata

The sixfold  $AT = (T, I(T), \vartheta, s_i, E, G)$  is referred to as a fully deterministic finite-state intuitionistic fuzzy automata (FDFIFA) in which  $s_i \in I(T)$  is the initial state;  $T$  is a finite family of states;  $I(T)$  is an intuitionistic fuzzy subset of states with grade of membership  $\mu: T \rightarrow (0, 1]$  and grade of nonmembership  $\nu: T \rightarrow [0, 1)$ ;  $E$  is a finite family of input letters/symbols;  $G$  is an intuitionistic fuzzy set of terminal/accepting states with  $G \subseteq I(T)$ ; and  $\vartheta: I(T) \times E \rightarrow I(T)$  is the transitional mapping with  $\vartheta((t_i, \mu^{t_i}, \nu^{t_i}), e) = (t_j, \mu^{t_j}, \nu^{t_j})$ , where  $(t_i, \mu^{t_i}, \nu^{t_i}), (t_j, \mu^{t_j}, \nu^{t_j}) \in I(T), e \in E$ . These states

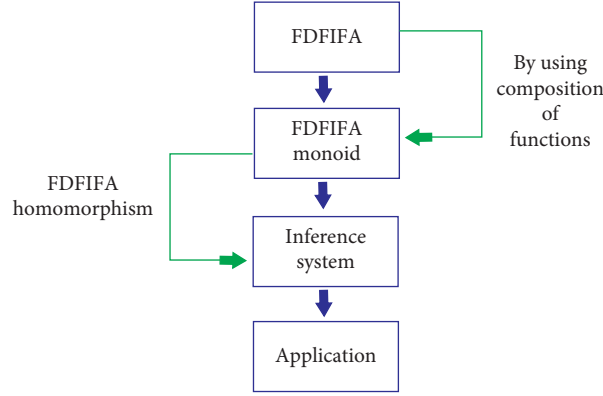


FIGURE 1: Fully intuitionistic fuzzy state transition.

$(t_j, \mu^{t_j}, \nu^{t_j}) \in I(T)$  exist in some accepted states. The extended transitional function  $\vartheta^*$  is defined as follows:

$$\vartheta^*((t_i, \mu^{t_i}, \nu^{t_i}), e_1 e_2 \dots e_n) = \begin{cases} (t_i, \mu^{t_i}, \nu^{t_i}), & \text{if } n = 0; \\ (\vartheta(\vartheta^*(t_i, \mu^{t_i}, \nu^{t_i}), e_1 e_2 \dots e_{n-1}), e_n), & \text{if } n \neq 0, \end{cases} \quad (1)$$

where  $e_1 e_2 \dots e_n \in E^*$  is the finite set of finite strings (words/sequences of input letters (symbols)) accepted by the automata.

*Example 1.* The following is considered:

- (1)  $T = \{t_1, t_2, t_3\}$ , the finite family of states.
- (2)  $I(T) = \{(t_i, \mu^{t_i}, \nu^{t_i}) : t_i \in T, \mu : T \rightarrow (0, 1], \nu : T \rightarrow (0, 1] \text{ and } 0 \leq \mu(t) + \nu(t) \leq 1\} = \{(t_1, 1, 0), (t_2, 0.4, 0.5), (t_3, 0.3, 0.7)\}$ , the finite family of states with  $s_i = (t_1, 1, 0)$  as initial state.
- (3)  $E = \{e_1, e_2\}$ , the collection of input letters/symbols.

The transitional mapping  $\vartheta : I(T) \times E \rightarrow I(T)$  for the accepting states  $t_3, t_2$  is defined as follows:

$$\begin{aligned} \vartheta((t_1, 1, 0), e_1) &= (t_1, 1, 0); \\ \vartheta((t_2, 0.4, 0.5), e_1) &= (t_1, 1, 0); \\ \vartheta((t_3, 0.3, 0.7), e_1) &= (t_2, 0.4, 0.5); \\ \vartheta((t_1, 1, 0), e_2) &= (t_3, 0.3, 0.7); \\ \vartheta((t_2, 0.4, 0.5), e_2) &= (t_3, 0.3, 0.7); \\ \vartheta((t_3, 0.3, 0.7), e_2) &= (t_3, 0.3, 0.7), \end{aligned} \quad (2)$$

with the transition diagram in Figure 2.

Then,  $G = \{(t_2, 0.4, 0.5), (t_3, 0.3, 0.7)\}$  is an intuitionistic fuzzy set of accepting states. Thus,  $AT = (T, I(T), \vartheta, s_i, E, G)$ . If we define mappings as follows:

$$\begin{aligned} f_{e_1}(t_1, 1, 0) &= \vartheta^*((t_1, 1, 0), e_1), \\ f_{e_1}(t_2, 0.4, 0.5) &= \vartheta^*((t_2, 0.4, 0.5), e_1), \\ f_{e_1}(t_3, 0.3, 0.7) &= \vartheta^*((t_3, 0.3, 0.7), e_1), \\ f_{e_2}(t_1, 1, 0) &= \vartheta^*((t_1, 1, 0), e_2), \\ f_{e_2}(t_2, 0.4, 0.5) &= \vartheta^*((t_2, 0.4, 0.5), e_2), \\ f_{e_2}(t_3, 0.3, 0.7) &= \vartheta^*((t_3, 0.3, 0.7), e_2), \\ f_{\wedge}(t_i, \mu^{t_i}, \nu^{t_i}) &= \vartheta^*((t_i, \mu^{t_i}, \nu^{t_i}), \wedge), \end{aligned} \quad (3)$$

where  $(t_i, \mu^{t_i}, \nu^{t_i}) \in I(T)$  and  $\wedge \in E^*$  (an empty sequence) make a table (see Table 1) for  $f_{\wedge}, f_{e_1}$  and  $f_{e_2}$  that shows their effects on the states  $t_1, t_2$  and  $t_3$ .

We are now interested to determine whether  $f_{\wedge}, f_{e_1}$  and  $f_{e_2}$  form a monoid under the composition of functions.

Elements  $f_{e_1 e_1}, f_{e_2 e_1}$  of Table 2 do not exist in the set. Therefore,  $\{f_{\wedge}, f_{e_1}, f_{e_2}\}$  is not a monoid. Now, it is expanded to  $\{f_{\wedge}, f_{e_1}, f_{e_2}, f_{e_1 e_1}, f_{e_2 e_1}\}$  and this is established as monoid  $M_{AT}$ , which corresponds to FDFIFA  $AT = (T, I(T), \vartheta, s_i, E, G)$ .

### 3. Monoid of Fully Deterministic Finite-State Intuitionistic Fuzzy Automata

Let  $AT = (T, I(T), \vartheta, s_i, E, G)$  be an FDFIFA. For any input string or sequence  $e_1 e_2 \dots e_n \in E^*$ , each state  $t_i \in T$  with

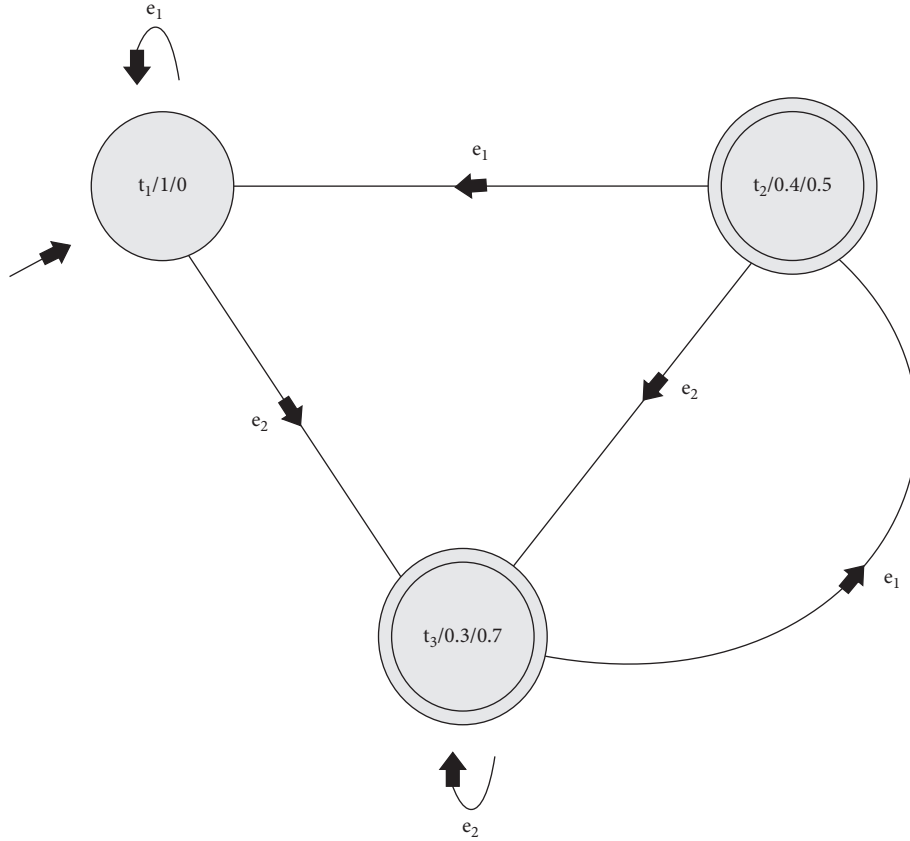


FIGURE 2: Fully intuitionistic fuzzy state transition.

TABLE 1: Mapping effects on states.

*	$f_{\wedge}$	$f_{e_1}$	$f_{e_2}$
$t_1$	$t_1$	$t_1$	$t_3$
$t_2$	$t_2$	$t_1$	$t_3$
$t_3$	$t_3$	$t_2$	$t_3$

TABLE 2: Composition of mappings.

$\circ$	$f_{\wedge}$	$f_{e_1}$	$f_{e_2}$
$f_{\wedge}$	$f_{\wedge}$	$f_{e_1}$	$f_{e_2}$
$f_{e_1}$	$f_{e_1}$	$f_{e_1 e_1}$	$f_{e_2 e_1}$
$f_{e_2}$	$f_{e_2}$	$f_{e_2}$	$f_{e_2}$

membership  $\mu^{t_i}$  and nonmembership  $\nu^{t_i}$  is transited to another state  $(t_j, \mu^{t_j}, \nu^{t_j}) \in G$ . If  $f_{e_1 e_2 \dots e_n}: I(T) \rightarrow I(T)$  is defined as  $f_{e_1 e_2 \dots e_n}(t_i, \mu^{t_i}, \nu^{t_i}) = \vartheta^*((t_i, \mu^{t_i}, \nu^{t_i}), e_1 e_2 \dots e_n)$  for each  $e_1 e_2 \dots e_n \in E^*$  and  $(t_i, \mu^{t_i}, \nu^{t_i}) \in I(T)$ , then  $M_{AT} = \{f_{e_1 e_2 \dots e_n}: e_1 e_2 \dots e_n \in E^*\}$  is a collection, which corresponds to FDFIFA AT. Clearly,  $\circ$  is a binary operation on  $M_{AT}$  such that for any  $f_{e_1 e_2 \dots e_k}, f_{e_1 e_2 \dots e_l}$  and  $f_{e_1 e_2 \dots e_m} \in M_{AT}$  with:

$$\begin{aligned}
 f_{e_1 e_2 \dots e_k}(t_i, \mu^{t_i}, \nu^{t_i}) &= \vartheta^*((t_i, \mu^{t_i}, \nu^{t_i}), e_1 e_2 \dots e_k), \\
 f_{e_1 e_2 \dots e_l}(t_i, \mu^{t_i}, \nu^{t_i}) &= \vartheta^*((t_i, \mu^{t_i}, \nu^{t_i}), e_1 e_2 \dots e_l), \\
 f_{e_1 e_2 \dots e_m}(t_i, \mu^{t_i}, \nu^{t_i}) &= \vartheta^*((t_i, \mu^{t_i}, \nu^{t_i}), e_1 e_2 \dots e_m),
 \end{aligned} \tag{4}$$

where  $(t_i, \mu^{t_i}, \nu^{t_i}) \in I(T)$  and  $k, l, m = 1, 2, \dots, n$  determine the length of string,  $I(T)$  is an intuitionistic fuzzy subset of states with grade of membership (HTML translation failed)

and grade of nonmembership  $\nu: T \rightarrow [0, 1]$ ,  $E^*$  is the finite set of finite strings (words/sequences of input letters/symbols) accepted by the automata, and  $\vartheta^*$  is the extended transitional function. We have the following:

$$f_{e_1 e_2 \dots e_k} \circ (f_{e_1 e_2 \dots e_1} \circ f_{e_1 e_2 \dots e_m}) = (f_{e_1 e_2 \dots e_k} \circ f_{e_1 e_2 \dots e_1}) \circ f_{e_1 e_2 \dots e_m}, \quad (5)$$

which turns FDFIFA  $M_{AT}$  into monoid, called an FDFIFA monoid (under the composition of mappings). The expression  $f_{e_1 e_2 \dots e_k} \circ f_{e_1 e_2 \dots e_1} = f_{e_1 e_2 \dots e_k e_1 e_2 \dots e_1}$  represents the overall state transition through  $\vartheta^*$  in the FDFIFA. FDFIFA contain a finite collection of states; thus, if the order of the collection

of  $I(T)$  is  $n$ , that is,  $|I(T)| = n$ , then the monoid order is  $|M_{AT}| \leq n^n$ .

*Definition 1.* Let  $(M_{AT}, \circ)$  be an FDFIFA monoid corresponding to the automata AT, where:

$$AT = (T, I(T), \vartheta, s_i, E, G), \vartheta((t_i, \mu^{t_i}, \nu^{t_i}), e) = (t_j, \mu^{t_j}, \nu^{t_j}). \quad (6)$$

The extended transitional function  $\vartheta^*$  is defined as follows:

$$\vartheta^*((t_i, \mu^{t_i}, \nu^{t_i}), e_1 e_2 \dots e_n) = \begin{cases} (t_i, \mu^{t_i}, \nu^{t_i}), & \text{if } n = 0; \\ (\vartheta(\vartheta^*(t_i, \mu^{t_i}, \nu^{t_i}), e_1 e_2 \dots e_{n-1}), e_n), & \text{if } n \neq 0. \end{cases} \quad (7)$$

$M_{AT1}$  is considered corresponding to the automata:

$$AT_1 = (T_1, I(T)_1, \vartheta_1, s_{i1}, E_1, G_1), \vartheta_1((t_{i1}, \mu^{t_{i1}}, \nu^{t_{i1}}), e) = (t_{j1}, \mu^{t_{j1}}, \nu^{t_{j1}}). \quad (8)$$

Then,  $\vartheta_1 = \vartheta/I(P)_1 \times E_1$ , and

$$\vartheta_1^* = \frac{\vartheta^*}{I(T)_1} \times E_1^*, \quad (9)$$

where  $(t_i, \mu^{t_i}, \nu^{t_i}), (t_j, \mu^{t_j}, \nu^{t_j}) \in I(T), e \in E, e_1 e_2 \dots e_n \in E^*, (t_{i1}, \mu^{t_{i1}}, \nu^{t_{i1}}), (t_{j1}, \mu^{t_{j1}}, \nu^{t_{j1}}) \in I(T)_1, e \in E_1 \subseteq E, T_1 \subseteq T, I(T)_1 \subseteq I(T), E_1^* \subseteq E^*, s_i = s_{i1}$  and  $G_1 \subseteq G$ . If  $M_{AT1}$  is closed under the binary operation  $\circ$ , symbolically,  $f_{e_1 e_2 \dots e_k} \circ f_{e_1 e_2 \dots e_1} \in M_{AT1} \forall f_{e_1 e_2 \dots e_k}, f_{e_1 e_2 \dots e_1} \in M_{AT1}$ , then  $(M_{AT1}, \circ)$  is known as a submonoid automaton of  $(M_{AT}, \circ)$ .

*Definition 2.* Any two FDFIFA monoids can be related to each other employing structure-preserving maps, commonly known as homomorphisms. Such maps not only preserve the binary operations used to design semigroups on FDFIFA but also the components involved in the formulation of FDFIFAs. Formally, if  $(M_{AT}, \circ)$  and  $(M_{AT1}, \circ_1)$  are any two FDFIFA monoids, then a map  $\theta: M_{AT} \rightarrow M_{AT1}$  defined as  $\theta(f_{e_1 e_2 \dots e_n}) = f_{\psi(e_1 e_2 \dots e_n)}$  is called an FDFIFA monoid homomorphism, and if  $\theta(f_{\wedge}) = f_{\psi(\wedge)} = f_{\wedge}$  and  $\forall f_{e_1 e_2 \dots e_k}, f_{e_1 e_2 \dots e_1} \in M_{AT}$ ,

$$\begin{aligned} \theta(f_{e_1 e_2 \dots e_k} \circ f_{e_1 e_2 \dots e_1}) &= \theta(f_{e_1 e_2 \dots e_k e_1 e_2 \dots e_1}) \\ &= f_{\psi(e_1 e_2 \dots e_k e_1 e_2 \dots e_1)} \\ &= f_{\psi(e_1 e_2 \dots e_k) \psi(e_1 e_2 \dots e_1)} \\ &= f_{\psi(e_1 e_2 \dots e_k)} \circ_1 f_{\psi(e_1 e_2 \dots e_1)} \\ &= \theta(f_{e_1 e_2 \dots e_k}) \circ_1 \theta(f_{e_1 e_2 \dots e_1}), \end{aligned} \quad (10)$$

where  $''\wedge''$  (an empty sequence)  $\in E^*$ ,  $''\wedge'''$  (an empty sequence)  $\in E_1^*$ , and  $\psi: E^* \rightarrow E_1^*$  is a mapping, defined as  $\forall e_1 e_2 \dots e_k, e_1 e_2 \dots e_l \in E^*, \psi(e_1 e_2 \dots e_k e_1 e_2 \dots e_l) = \psi(e_1 e_2 \dots e_k) \psi(e_1 e_2 \dots e_l)$ . If  $\theta$  is a homomorphism that also establishes a one-to-one correspondence between  $M_{AT}$  and  $M_{AT1}$ , then it is termed to be an isomorphism and the FDFIFA monoids are denoted by  $M_{AT} \cong M_{AT1}$ .

#### 4. Intuitionistic Fuzzy Inference Rule on FDFIFA Monoid

On the semigroup of FDFIFA, intuitionistic fuzzy inference is a cognitive procedure that proceeds to a new decision or statement based on one or more previous decisions or propositions. In general, FDFIFA inference is divided into two components. The premise is a well-known decision used as a starting point for inference. The second is the conclusion, which is a fresh decision generated by the proposition. The following inference rules are offered based on the characteristics of homomorphic mapping:

Premise 1: if  $a \in X$  semigroup, then  $b \in Y$  must be a semigroup.

Premise 2: if  $a \in X$  is a semigroup, then

Conclusion:  $b \in Y = X \circ (X \rightarrow Y)$ , where  $\circ$  is a binary operation and  $X \rightarrow Y$  is a homomorphism.

The intuitionistic fuzzy inference system is made up of certain inference rules that must obey certain arithmetic requirements. The relationship between  $X$  and  $Y$  for Premise 1 is  $X \rightarrow Y$ , which is a homomorphism, grades of membership and nonmembership corresponding to an intuitionistic fuzzy relation matrix  $F$ , which is defined as follows:

$$\begin{aligned}
F(p, q) &= (\mu_{X \rightarrow Y}(p, q), \nu_{X \rightarrow Y}(p, q)) \\
&= [(\mu_X(p) \wedge \mu_Y(q)) \vee (1 - \mu_X(p)), (\nu_X(p) \vee \nu_Y(q)) \wedge (1 - \nu_X(p) - \epsilon)],
\end{aligned} \tag{11}$$

where  $\mu_{X \rightarrow Y}(p, q), \nu_{X \rightarrow Y}(p, q)$  are the grades of membership and nonmembership for intuitionistic inference rule,  $\epsilon$  is an error of hesitancy,  $\mu_Y(q), \nu_Y(q)$  are the grades of membership and nonmembership for the assumption that  $q \in Y$  is a semigroup, and  $\mu_X(p), \nu_X(p)$  are the grades of membership and nonmembership for the assumption that  $p \in X$  is a semigroup, and  $\vee$  is an “or” operation and  $\wedge$  is an “and” operation:

$$Y = X \circ (X \rightarrow Y). \tag{12}$$

The inference relationship between  $X$  and  $X$  to  $Y$  can be used to synthesize the conclusion  $Y$ .  $F$  can be used to obtain the conclusion’s membership and nonmembership functions:

$$(\mu_Y(q), \nu_Y(q)) = (\mu_X(p), \nu_X(p)) \circ F(p, q). \tag{13}$$

The following is a description of  $\circ$ , the synthetic binary operator:

$$\begin{aligned}
(\mu_Y(q), \nu_Y(q)) &= (\mu_X(p), \nu_X(p)) \circ F(p, q) \\
&= \left[ \bigvee_{p \in X} \left\{ \mu_X(p) \wedge \mu_F(p, q) \right\}, \bigwedge_{p \in X} \left\{ \nu_X(p) \vee \nu_F(p, q) \right\} \right],
\end{aligned} \tag{14}$$

such that  $0 \leq \mu_Y(q) + \nu_Y(q) \leq 1$ . In the classical fuzzy inference system, grade on membership of the fuzzy subset is a variable and variables are related through the operations defined on membership grades. However, in the inference system of FDFIFA semigroup, the FDFIFA semigroup serves as a variable and relationship between variables is exhibited using FDFIFA semigroup homomorphism.

## 5. Application

Automobiles provide millions around the world with a sense of freedom. People may now live, work, and engage in ways that were previously impossible a century ago. The automobile industry is the world’s largest single most powerful economic driver. The earliest steam-powered vehicle was invented in 1672, and Nicolas-Joseph Cugnot created the first steam-powered automotive capable of human movement in 1770 [37], since then several automotive models are introduced with numerous modifications [38, 39]. The consumption of nonrenewable fuels, a major increase in the likelihood of unintentional fatality, the emission of noise and air pollution, and the production of greenhouse gases are only a few of the present negative repercussions of widespread automobile use. The key benefit of a hybrid automobile is that it uses less gasoline and emits less CO<sub>2</sub> [40] than a typical petrol or diesel-powered vehicle.

A hybrid vehicle is one that has multiple modes of propulsion, such as a gasoline or diesel engine and an electric motor. There are several kinds of hybrids, each of which functions differently [41]. Range extender hybrid automobiles (series hybrids) rely on their conventional engine to generate electricity to power a generator that recharges the batteries. Instead of propelling the car, the engine provides energy to the electric motor [42]. The well-known model is BMW i3 [43]. Plug-in hybrids are a kind of hybrid that

can be recharged both at home and on the go, as the title indicates [44]. The BMW 330e, Mitsubishi Outlander, and Volvo V60 are just a few of the plug-in hybrids on the market. The hybrid car moves by combining at least one electrical motor with a gasoline engine, while the mechanism recovers energy through regenerative braking [45]. As result, lesser gasoline is burnt, resulting in higher fuel economy. Most of them use a high-voltage battery pack (different from the car’s standard 12-volt battery) that is recharged by absorbing energy from deceleration that would otherwise be wasted to heat created by the brakes in traditional cars. (This is accomplished through the use of regenerative braking.)

The basic purpose of such regenerative braking is to transform kinetic energy (from the wheels) towards electrical energy and save it in batteries for future use in vehicle propulsion [46]. For reliable vehicle operation and control, these systems often use electric motors for both traction and regenerative braking. They also entail constant interactions between mechanical/electrical components and their controllers (electronic control units, ECUs). Faults in these technologies can have a substantial impact on the reliability and efficiency of automobile control and operation. According to the available literature, the focus of hybrid electric vehicle (HEV) studies has been on design and control methods for enhanced energy regeneration [47], best possible braking strategies [48], and mechanism level control to improve energy/fuel efficiency [49–51], with fault analysis receiving far less attention. A fault-tolerant power train topology for series hybrid electric vehicles was discussed by Song and Wang [52]. They looked at short-switch and open-switch failures that threaten motor drive unit reliability. Parsa and Toliyat [53] suggested a fault-tolerant control approach for five-phase permanent magnet motors. It was determined that the system could function securely with up to two phases lost without the need of any extra



equipment. A basic on-board fault detection technique built on reference frame theory for detecting electric motor defects in HEVs at start-up and idle circumstances is described in [54]. Rothenhagen and Fuchs [55] present a spectator residual generation technique to identify current sensor problems.

In cascaded multiple converter devices, Jayabalan and Fahimi [56] proposed using statistical moments of higher powers to identify open-circuit and short-circuit failures. Merzouki et al. [57] discussed a parity relation-based residual generation technique for detecting and isolating actuator faults in electric vehicles. Existing HEV failure diagnostic methods are component-centric and may not explicitly incorporate communication or scheme interactions. This research looks at hardware, programming, and communication defects in HEV's regenerative braking (RBS) and proposes a data-driven technique for detecting and identifying them. Parametric and sensor-related problems (underlying physical faults), software logic faults, and interprocess communication defects (missed messages, several messages, and obsolete message faults) are among the defects investigated. The use of a series-parallel power train with regenerative braking is contemplated. The automobile is normally propelled by two power sources: (i) an internal combustion engine (ICE) and an electric generator; and (ii) an electric motor having a battery as the power storage system [58]. The various constituent components of the drive train design (with a series-parallel power train

configuration) are shown in Figure 3. The car operates in an electricity-only mode when the speed or power demand is minimal (series mode). When the car's power requirement at the wheels is larger, the engine and the motor work together to push the vehicle forward (parallel mode). Ehsani et al. [59] provided further information about different drive train layouts. As a result, defining nominal and problematic behaviour, as well as developing detection and inference algorithms for rapid fault diagnosis, is crucial.

**5.1. FDFIFA Monoid.** Now, we construct an FDFIFA semigroup  $(M_{AT}, \circ)$  fault diagnosis model for the hybrid electric vehicle regenerative braking system. Let  $(M_{AT}, \circ)$  be an FDFIFA semigroup corresponding to the automata  $AT = (T, I(T), \vartheta, s_i, E, G)$ , where  $T$  is a collection of states that the hybrid electric vehicle regenerative braking system can be in, such as normal, medium, or serious faults;  $I(T)$  is an intuitionistic fuzzy subset of states  $T$  with grade of membership  $\mu: T \rightarrow (0, 1]$  and grade of nonmembership  $\nu: T \rightarrow [0, 1)$ ;  $E$  is an input character signal set;  $s_i \in I(T)$  is the initial state of the processing signal;  $G$  is a fault state set and a subset of  $I(T)$ ; and  $\vartheta: I(T) \times E \rightarrow I(T)$  is the transitional mapping with  $\vartheta((t_i, \mu^{t_i}, \nu^{t_i}), e) = (t_j, \mu^{t_j}, \nu^{t_j})$ , where  $(t_i, \mu^{t_i}, \nu^{t_i}), (t_j, \mu^{t_j}, \nu^{t_j}) \in I(T), e \in E$ . These states  $(t_j, \mu^{t_j}, \nu^{t_j}) \in I(T)$  exist in some accepted states. The extended transitional function  $\vartheta^*$  is defined as follows:

$$\vartheta^*((t_i, \mu^{t_i}, \nu^{t_i}), e_1 e_2 \dots e_n) = \begin{cases} (t_i, \mu^{t_i}, \nu^{t_i}), & \text{if } n = 0; \\ (\vartheta(\vartheta^*(t_i, \mu^{t_i}, \nu^{t_i}), e_1 e_2 \dots e_{n-1}), e_n), & \text{if } n \neq 0, \end{cases} \quad (15)$$

where  $e_1 e_2 \dots e_n \in E^*$  is the finite set of finite strings (words/sequences of input letters (symbols)) accepted by the automata. Because  $s_i \in I(T)$ ,  $I(T)$  denotes the entire states,  $G \subseteq I(T)$ , and  $\vartheta$  is a transition procedure from one state to another. The degree of FDFIFA semigroup membership and nonmembership can show the regenerative braking's degree of fault and accuracy. As a result, the degree of membership is utilized to indicate the severity of the defect.  $G$  and  $E$  are mostly treated in the following order:

- (i) Fault state or output variable set **G**: the normal working condition represented by  $W_1$ , battery current sensor fault represented by  $W_2$ , battery temperature sensor fault denoted by  $W_3$ , engine speed sensor fault denoted by  $W_4$ , motor 1 current sensor fault denoted by  $W_5$ , motor 1 speed sensor fault denoted by  $W_6$ , vehicle speed sensor fault denoted by  $W_7$ , wheel inertia fault denoted by  $W_8$ , engine message loss fault denoted by  $W_9$ , burst loss of engine message denoted by  $W_{10}$ , burst loss of message from PTC denoted by  $W_{11}$ , too many messages from battery denoted by  $W_{12}$ , battery initial SOC fault denoted by  $W_{13}$ , wheel radius fault denoted by  $W_{14}$ , engine message faulty data denoted

by  $W_{15}$ , motor 1 message faulty data denoted by  $W_{16}$ , and wheel message faulty data denoted by  $W_{17}$ . The range of values from  $W_1$  to  $W_{17}$  is included inside the  $[0, 1]$  interval. 0 in the interval denotes the absence of such a flaw. 1 denotes a major flaw [60].

- (ii) Processing signal or input variable set **E**: the following 25 felt parameters are chosen as input variables:  $e_1$ : battery SOC,  $e_2$ : motor 2 torque demand,  $e_3$ : wheel torque demand,  $e_4$ : motor 1 torque demand,  $e_5$ : engine torque demand,  $e_6$ : battery temperature,  $e_7$ : battery current,  $e_8$ : driver torque demand,  $e_9$ : motor 1 command,  $e_{10}$ : gearbox speed,  $e_{11}$ : wheel input speed,  $e_{12}$ : wheel output speed,  $e_{13}$ : wheel torque,  $e_{14}$ : vehicle linear speed,  $e_{15}$ : motor 1 speed,  $e_{16}$ : motor 1 current,  $e_{17}$ : clutch input speed,  $e_{18}$ : engine command,  $e_{19}$ : motor 2 command,  $e_{20}$ : motor 2 speed,  $e_{21}$ : motor 2 current,  $e_{22}$ : engine speed,  $e_{23}$ : clutch output speed,  $e_{24}$ : mechanical accessory torque, and  $e_{25}$ : wheel command [60].
- (iii) The degrees to which each fault and accuracy parameter corresponds to the categories of "normal," "serious," and "medium" are presented in terms of

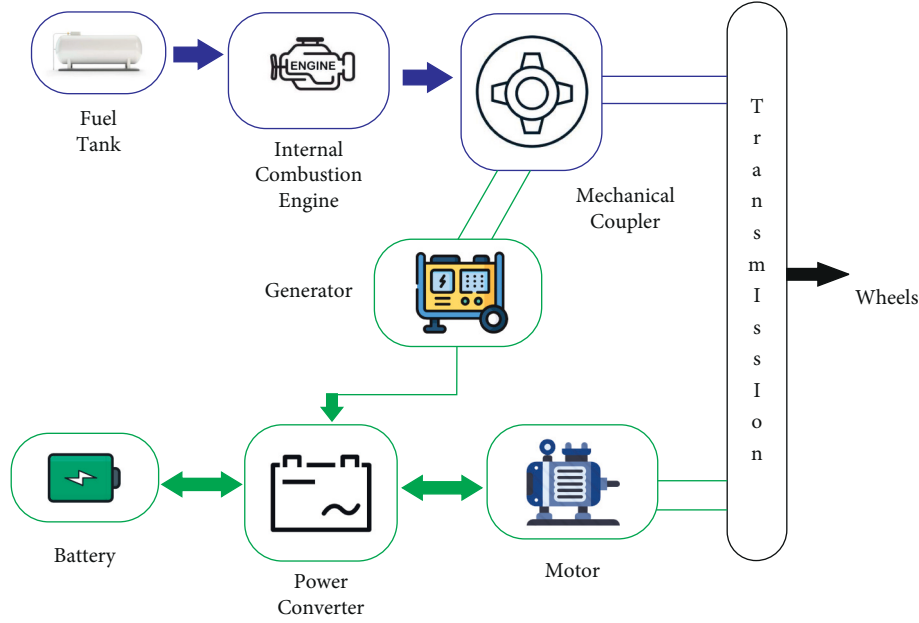


FIGURE 3: Series-parallel power train.

membership and the intuitionistic degree of non-membership grades as follows:

$$(\mu_{\text{normal}}, \nu_{\text{normal}}) = \begin{pmatrix} (1, 0) (0.2, 0.6) (0.1, 0.8) (0.2, 0.5) (0.1, 0.8) \\ (0, 0.9) (0, 0.8) (0.1, 0.6) (0.1, 0.7) \\ (0, 0.4) (0.2, 0.6) (0.1, 0.8) (0, 0.9) (0.2, 0.6) (0.2, 0.7) (0, 0.9) (0.1, 0.8) \end{pmatrix}, \quad (16)$$

$$(\mu_{\text{serious}}, \nu_{\text{serious}}) = \begin{pmatrix} (0, 0.6) (0.2, 0.7) (0.3, 0.7) (0.4, 0.5) (0.6, 0.2) (0.8, 0.1) (0.6, 0.3) \\ (0.7, 0.2) (0.1, 0.7) (0.3, 0.6) \\ (0.6, 0.4) (0.2, 0.8) (0.1, 0.9) (0.7, 0.2) (0.3, 0.7) (0.1, 0.9) (0.2, 0.8) \end{pmatrix}, \quad (17)$$

$$(\mu_{\text{medium}}, \nu_{\text{medium}}) = \begin{pmatrix} (1, 0) (0.3, 0.4) (0.2, 0.8) (0.1, 0.9) (0.4, 0.3) (0.3, 0.4) (0.3, 0.7) \\ (0.2, 0.7) (0.1, 0.8) (0, 0.6) (0, 0.9) (0.5, 0.4) \\ (0.2, 0.7) (0.3, 0.2) (0.4, 0.6) (0.2, 0.5) (0.4, 0.5) \end{pmatrix}. \quad (18)$$

**5.2. Intuitionistic Fuzzy Inference Model.** Faults are divided into three categories to make diagnosis easier: serious fault, medium fault, and no fault. The result is processed as follows, based on the output state of the intuitionistic fuzzy inference model of FDFIFA semigroup:

If  $0.60 \leq \mu(W_i) \leq 1$  and  $0 \leq \nu(W_i) \leq 0.40$ ,  $W_i$  is level 1, recognized as the serious fault.

When  $0.25 < \mu(W_i) < 0.60$  and  $0.40 < \nu(W_i) < 0.75$ ,  $W_i$  is level 2, recognized as the medium defect.

If  $\mu(W_i) \leq 0.25$  and  $\nu(W_i) \leq 0.75$ ,  $W_i$  is normal, recognized as no fault.

The seriousness of the defect can be obtained by utilizing this information. The FDFIFA semigroup's inference system has now been completely specified, comprising variables, membership functions, nonmembership functions, and the essential rules for diagnosing faults. The FDFIFA semigroup

rule viewer permits us to totally comprehend the entire intuitionistic fuzzy inference process at once. It also demonstrates the influences of the membership and non-membership functions on the overall outcomes of the intuitionistic fuzzy inference. Our prior work has a full description of how to design an intuitionist fuzzy inference system for fault detection [24]. If the problem being examined is complex in nature subject to several conflicting factors, then FDFIFA semigroup inference is an appropriate tool for its solution. Semigroup's intuitionistic fuzzy inference model is put to the test with a collection of real data to see whether it can correctly identify a hybrid electric vehicle regenerative braking system issue. The intuitionistic fuzzy inference model can be employed if the fault can be appropriately diagnosed. The corresponding parameters are monitored and their values are acquired when  $W_5, W_6, W_7,$



and  $W_8$  are defective. The data are utilized as input vectors in simulation functions to compute the output of an intuitionistic fuzzy inference model. The fault criterion is used to

determine whether or not a flaw exists. The test result shows that 5, 6, and 12 in  $W_i - out$  signify distinct types of faults illustrated as follows:

---


$$\begin{aligned} \text{Input vector} &= \begin{pmatrix} (0.2, 0.7) (0.1, 0.7) (0.2, 0.7) (0.6, 0.3) (0, 0.8) \\ (0, 0.9) (0.1, 0.7) (0.1, 0.8) (0, 0.9) (0.1, 0.7) (0.2, 0.7) \\ (0, 0.8) (0, 0.9) (0.1, 0.7) (0.1, 0.8) (0, 0.9) \end{pmatrix}, \\ \text{Cause of fault} &= W_5 \text{ serious fault,} \\ \text{Input vector} &= \begin{pmatrix} (0.2, 0.6) (0.1, 0.7) (0.2, 0.6) (0.1, 0.8) (0.4, 0.5) \\ (0, 0.8) (0.1, 0.7) (0.1, 0.8) (0, 0.9) (0.2, 0.6) (0.1, 0.7) \\ (0.2, 0.6) (0.1, 0.7) (0.2, 0.6) (0.1, 0.7) (0.2, 0.6) \end{pmatrix}, \\ \text{Cause of fault} &= W_6 \text{ medium fault,} \\ \text{Input vector} &= \begin{pmatrix} (0.2, 0.6) (0.1, 0.8) (0.2, 0.6) (0.1, 0.7) (0.2, 0.6) \\ (0.1, 0.7) (0.2, 0.6) (0.2, 0.7) (0.1, 0.8) (0, 0.8) (0.6, 0.3) \\ (0.1, 0.8) (0.1, 0.7) (0, 0.8) (0.1, 0.8) (0.1, 0.7) \end{pmatrix}, \\ \text{Cause of fault} &= W_{12} \text{ serious fault.} \end{aligned} \tag{19}$$

Suppose  $(M_{AT1}, \circ_1)$  be an FDFIFA semigroup corresponding to the automata  $AT1$ ,  $(M_{AT2}, \circ_2)$  be an FDFIFA semigroup corresponding to the automata  $AT2$  as constructed above, and  $\theta: M_{AT1} \rightarrow M_{AT2}$  is an inference rule in terms of FDFIFA homomorphism. In the aforesaid application, the inference model for fault diagnostics can be built as follows:

Premise 1: if  $M_{AT1}$  parameter  $a$  is normal,  $M_{AT2}$  output  $b$  is serious.

Premise 2: if  $M_{AT1}$  parameter  $a$  is medium, then

---

Conclusion: try to figure out what the fault level of the output  $b$  in  $M_{AT2}$ .

The degree of membership and nonmembership that corresponds to each inference step of the model can be determined using the intuitionistic fuzzy inference system, starting with the known condition.

Step 1: the intuitionistic fuzzy relation matrix  $F$  can be calculated using Premise 1 and the formula (11), as well as the intuitionistic fuzzy degree of membership and nonmembership given by the aforementioned expressions (16) and (17).



$$\begin{aligned}
\left( \mu_{M_{AT2}}(\hat{b}), \nu_{M_{AT2}}(\hat{b}) \right) &= \left( \mu_{M_{AT1}}(\hat{a}), \nu_{M_{AT1}}(\hat{a}) \right) \circ F(a, b) \\
&= \begin{pmatrix} (1, 0) (0.3, 0.4) (0.2, 0.8) (0.1, 0.9) (0.4, 0.3) (0.3, 0.4) \\ (0.3, 0.7) (0.2, 0.7) (0.1, 0.8) (0, 0.6) (0, 0.9) (0.5, 0.4) \\ (0.2, 0.7) (0.3, 0.2) (0.4, 0.6) (0.2, 0.5) (0.4, 0.5) \circ F(a, b) \\ (0.5, 0.3) (0.5, 0.3) (0.5, 0.3) (0.5, 0.3) (0.6, 0.2) (0.8, 0.1) \\ (0.6, 0.3) (0.7, 0.2) (0.5, 0.3) (0.5, 0.3) (0.6, 0.3) (0.5, 0.3) \\ (0.5, 0.3) (0.7, 0.2) (0.5, 0.3) (0.5, 0.3) (0.5, 0.3) \end{pmatrix}. \tag{21}
\end{aligned}$$

When compared to the degree  $(\mu_{\text{serious}}, \nu_{\text{serious}})$

$$\begin{pmatrix} (0, 0.6) (0.2, 0.7) (0.3, 0.7) (0.4, 0.5) (0.6, 0.2) (0.8, 0.1) (0.6, 0.3) (0.7, 0.2) \\ (0.1, 0.7) (0.3, 0.6) (0.6, 0.4) (0.2, 0.8) (0.1, 0.9) (0.7, 0.2) (0.3, 0.7) (0.1, 0.9) (0.2, 0.8) \end{pmatrix}. \tag{22}$$

of “serious” in  $M_{AT2}$ , which corresponds to the fault parameters  $W_1, W_2, \dots, W_{17}$ , the degree of  $b$  in  $M_{AT2}$  is as follows:

$$\begin{pmatrix} (0.5, 0.3) (0.5, 0.3) (0.5, 0.3) (0.5, 0.3) (0.6, 0.2) (0.8, 0.1) (0.6, 0.3) (0.7, 0.2) \\ (0.5, 0.3) (0.5, 0.3) (0.6, 0.3) (0.5, 0.3) (0.5, 0.3) (0.7, 0.2) (0.5, 0.3) (0.5, 0.3) (0.5, 0.3) \end{pmatrix}, \tag{23}$$

which signifies “more serious.” If  $\hat{a}$  is a medium fault, the inference result for the output  $b \in M_{AT2}$  is a more serious fault, which corresponds to objective reality.

## 6. Conclusion

Traditional HEV failure diagnostic methods are component-centric and may not explicitly incorporate communication or scheme interactions. This research assessed hardware, programming, and communication flaws in a hybrid electric vehicle’s regenerative braking system (RBS) and suggested a data-driven method for discovering and identifying them. The defects investigated include parametric and sensor-related issues (underlying physical faults), software logic flaws, and interprocess communication flaws (missed messages, multiple messages, and obsolete message faults).

The intuitionistic fuzzy set (IFS) is a well-known generalization of fuzzy sets that have been thoroughly studied, with significant research into its theocratic properties and applications in various disciplines. Because nonmembership grades are involved, IFS can handle uncertainties better than fuzzy sets. In this research, the effectiveness of IFS is used to detect flaws in a regenerative braking system. First, the finite deterministic fully intuitionistic fuzzy automata (FDFIFA) are defined, and then, a semigroup over FDFIFA is designed.

The algebraic properties of the FDFIFA semigroup are investigated and used to construct inference systems over it. The proposed method outperforms the previously published fuzzy inference method. Fuzzy inference can be obtained from intuitionistic fuzzy inference by considering only the membership values.

Furthermore, any machine’s maximum default and present running states can be diagnosed using the maximum value of the membership and nonmembership grades. Based on the given target qualities, the current techniques must create equations or expressions to deal with the target. The described approach, on the other hand, uses its state transition mapping to execute characteristic processing and just requires the selection of system parameters. There are a number of practical generalizations of fuzzy sets [9, 61–68], and inference rules for these generalizations can be developed using the same techniques. The proposed methodology’s superiority is demonstrated by the real-life example, which accurately and efficiently discovers the defect. The proposed methodology can be used to detect a failure in any machine by identifying its critical components, the factors that can cause any fault in those components, and the fault caused by these parameters. The nature and severity of the defect can be used to award membership and nonmembership grades.

## Data Availability

Data are collected from relevant stakeholders through questionnaire.

## Conflicts of Interest

The authors declare that there were no conflicts of interest regarding the publication of this article.

## Authors' Contributions

All authors contributed equally to the preparation of this manuscript.

## References

- [1] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, no. 3, pp. 338–353, 1965.
- [2] K. Atanassov, "Intuitionistic fuzzy sets," *International Journal Bioautomation*, vol. 20, no. 1, pp. 87–96, 2016.
- [3] Y. Cao and Y. Ezawa, "Nondeterministic fuzzy automata," *Information Sciences*, vol. 191, pp. 86–97, 2012.
- [4] S. Singh and S. P. Tiwari, "On unification of categories of fuzzy automata as Qua category," *Soft Computing*, vol. 26, no. 4, pp. 1509–1529, 2022.
- [5] O. Castillo, P. Melin, and J. Kacprzyk, *Intuitionistic and Type-2 Fuzzy Logic Enhancements in Neural and Optimization Algorithms: Theory and Applications*, Springer, Berlin, Germany, 2020.
- [6] V. Karthikeyan, N. Mohanarao, and S. Sivamani, "Generalized products of directable fuzzy automata," *Materials Today Proceedings*, vol. 37, pp. 3531–3533, 2021.
- [7] M. Shamsizadeh and M. M. Zahedi, "Intuitionistic general fuzzy automata," *Soft Computing*, vol. 20, no. 9, pp. 3505–3519, 2016.
- [8] A. P. Singha, M. Dubeyb, and I. Perfilieva, "On quotient structures of fuzzy multiset finite automata," in *Proceedings of the Joint 19th World Congress of the International Fuzzy Systems Association (IFSA)*, Daegu, Korea, August 2021.
- [9] Y. Wang and L. Zhang, "Algebraic properties of intuitionistic fuzzy multiset finite automata," *Soft Computing*, vol. 25, no. 22, pp. 13857–13867, 2021.
- [10] Q. Wu, X. Wang, Z. Chen, H. Chen, D. Sun, and Z. Han, "Application of fuzzy pushdown automaton on prediction of quality control for spinning yarn," *Information Technology and Control*, vol. 50, no. 1, pp. 76–88, 2021.
- [11] Z. Jančić, I. Micić, J. Ignjatović, and M. Čirić, "Further improvements of determinization methods for fuzzy finite automata," *Fuzzy Sets and Systems*, vol. 301, pp. 79–102, 2016.
- [12] Q. E. Wu, T. Wang, X. Pang, Y. X. Huang, and J. S. Li, "Discussion of relation of fuzzy automata," *Journal of Information and Decision Science*, vol. 2, no. 4, pp. 349–357, 2007.
- [13] L. Wei, G. Huang, S. H. Wasti, M. J. Hussain, and Y. Jiang, "Generalized fuzzy automata with semantic computing," *Soft Computing*, vol. 25, no. 7, pp. 5775–5789, 2021.
- [14] S. Yadav, S. P. Tiwari, M. Kumari, and V. K. Yadav, "An interval type-2 fuzzy model of computing with words via interval type-2 fuzzy finite rough automata with application in covid-19 deduction," *New Mathematics and Natural Computation*, vol. 18, no. 1, pp. 61–101, 2022.
- [15] L. Wending, "Implementing the hybrid neuro-fuzzy system to model specific learning disability in special university education programs," *Journal of Mathematics*, vol. 2022, Article ID 6540542, 7 pages, 2022.
- [16] R. Marzoug, N. Lakouari, H. Ez-Zahraouy, B. Castillo Téllez, M. Castillo Téllez, and L. Cisneros Villalobos, "Modeling and simulation of car accidents at a signalized intersection using cellular automata," *Physica A: Statistical Mechanics and its Applications*, vol. 589, Article ID 126599, 2022.
- [17] C.-y. Li, S.-k. Liu, G.-m. Xu, and X.-k. Cen, "Influence of driver's yielding behavior on pedestrian-vehicle conflicts at a two-lane roundabout using fuzzy cellular automata," *Journal of Central South University*, vol. 29, no. 1, pp. 346–358, 2022.
- [18] L. Li and D. Qiu, "On the state minimization of fuzzy automata," *IEEE Transactions on Fuzzy Systems*, vol. 23, no. 2, pp. 434–443, 2014.
- [19] S. Moghari and M. M. Zahedi, "Minimization of deterministic fuzzy tree automata," *Journal of Fuzzy Set Valued Analysis*, vol. 2014, pp. 1–18, 2014.
- [20] J. R. González de Mendivil and J. R. Garitagoitia, "Determinization of fuzzy automata via factorization of fuzzy states," *Information Sciences*, vol. 283, pp. 165–179, 2014.
- [21] A. Stamenkovic, M. Ciric, and J. Ignjatovic, "Reduction of fuzzy automata by means of fuzzy quasi-orders," *Information Sciences*, vol. 275, pp. 168–198, 2014.
- [22] K. Krohn and J. Rhodes, "Algebraic theory of machines. I. prime decomposition theorem for finite semigroups and machines," *Transactions of the American Mathematical Society*, vol. 116, pp. 450–464, 1965.
- [23] M. M. K. Rao, "Tripolar fuzzy interior ideals and tripolar fuzzy soft interior ideals over semigroups," *Annals of Fuzzy Mathematics and Informatics*, vol. 20, no. 3, pp. 243–256, 2020.
- [24] S. Kousar, F. Aslam, N. Kausar, and Y. U. Gaba, "Semigroup of finite-state deterministic intuitionistic fuzzy automata with application in fault diagnosis of an aircraft twin-spool turbofan engine," *Journal of Function Spaces*, vol. 2021, Article ID 1994732, 10 pages, 2021.
- [25] R. Hampel, M. Wagenknecht, and N. Chaker, *Fuzzy Control: Theory and Practice*, Springer Science & Business Media, Berlin, Germany, 2013.
- [26] R. Tabbussum and A. Q. Dar, "Performance evaluation of artificial intelligence paradigms-artificial neural networks, fuzzy logic, and adaptive neuro-fuzzy inference system for flood prediction," *Environmental Science and Pollution Research*, vol. 28, no. 20, pp. 25265–25282, 2021.
- [27] V. Gholami, M. R. Khaleghi, S. Pirasteh, and M. J. Booi, "Comparison of self-organizing map, artificial neural network, and co-active neuro-fuzzy inference system methods in simulating groundwater quality: geospatial artificial intelligence," *Water Resources Management*, vol. 36, no. 2, pp. 451–469, 2022.
- [28] S. K. Paul, P. Chowdhury, K. Ahsan, S. M. Ali, and G. Kabir, "An advanced decision-making model for evaluating manufacturing plant locations using fuzzy inference system," *Expert Systems with Applications*, vol. 191, Article ID 116378, 2022.
- [29] H. H. H. Aly, "A hybrid optimized model of adaptive neuro-fuzzy inference system, recurrent Kalman filter and neuro-wavelet for wind power forecasting driven by DFIG," *Energy*, vol. 239, Article ID 122367, 2022.
- [30] S. K. Das, A. Das, H. K. Sinha, S. P. Das, and S. Nayak, "Game theory based optimal decision-making system," *Smart Innovation, Systems and Technologies*, vol. 194, pp. 761–769, Springer, Singapore, 2021.

- [31] I. B. Turksen and D. D. W. Yao, "Representations of connectives in fuzzy reasoning: The view through normal forms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 14, no. 1, pp. 146–151, 1984.
- [32] W. Zeng and H. Li, "Relationship between similarity measure and entropy of interval valued fuzzy sets," *Fuzzy Sets and Systems*, vol. 157, no. 11, pp. 1477–1484, 2006.
- [33] J. R. G. de Mendivil, "Conditions for minimal fuzzy deterministic finite automata via Brzozowski's procedure," *IEEE Transactions on Fuzzy Systems*, vol. 26, no. 4, pp. 2409–2420, 2017.
- [34] Z. Gao, Q. Wu, H. Chen, X. Qian, Y. Wang, and H. Zhang, "Target classification by constructing fuzzy automata system," *International Journal of Fuzzy Systems*, vol. 20, no. 8, pp. 2620–2631, 2018.
- [35] Y. Du and P. Zhu, "Fuzzy approximations of fuzzy relational structures," *International Journal of Approximate Reasoning*, vol. 98, pp. 1–10, 2018.
- [36] W. G. Wee and K. S. Fu, "A formulation of fuzzy automata and its application as a model of learning systems," *Readings in Fuzzy Sets for Intelligent Systems*, vol. 5, pp. 892–900, 1993.
- [37] O. Chamberlain and E. Ede, "History of automobile past and present challenges facing automobile production in Nigeria," *IOSR Journal of Research & Method in Education*, vol. 2, no. 4, pp. 11–16, 2013.
- [38] S. Balamurugan, "Evaluation of heat treatment characteristics for case hardening steels in automobiles," in *Proceedings of the 32 nd Indian Engineering Congress, The Institution of Engineers, Chennai, India, November 2021*.
- [39] A. V. Terentyev, M. Y. Karelina, T. Y. Cherepnina, D. A. Linnik, and V. A. Demin, "Digital object-oriented control models in automobile-road complex systems," *IOP Conference Series: Materials Science and Engineering*, vol. 832, no. 1, Article ID 012058, 2020.
- [40] P. Plötz, S. Á. Funke, and P. Jochem, "Empirical fuel consumption and CO<sub>2</sub> emissions of plug-in hybrid electric vehicles," *Journal of Industrial Ecology*, vol. 22, no. 4, pp. 773–784, 2018.
- [41] R. Bejgam, S. Sunkari, S. B. Keshipeddi, M. R. Rangaraju, and V. Dunde, "A brief study on hybrid electric vehicles," in *Proceedings of the 2021 3rd International Conference on Inventive Research in Computing Applications (ICIRCA)*, pp. 54–59, Coimbatore, India, September 2021.
- [42] M. I. Tametang Meli, D. Yemélé, and G. D. Leutcho, "Dynamical analysis of series hybrid electric vehicle powertrain with torsional vibration: Antimonotonicity and coexisting attractors," *Chaos, Solitons & Fractals*, vol. 150, Article ID 111174, 2021.
- [43] M. Gierczynski and L. M. Grzesiak, "Comparative analysis of the steady-state model including non-linear flux linkage surfaces and the simplified linearized model when applied to a highly-saturated permanent magnet synchronous machine-evaluation based on the example of the bmw i3 traction motor," *Energies*, vol. 14, no. 9, p. 2343, 2021.
- [44] A.-M. Hariri, M. A. Hejazi, and H. Hashemi-Dezaki, "Investigation of impacts of plug-in hybrid electric vehicles' stochastic characteristics modeling on smart grid reliability under different charging scenarios," *Journal of Cleaner Production*, vol. 287, Article ID 125500, 2021.
- [45] S. R. Cikanek and K. E. Bailey, "Regenerative braking system for a hybrid electric vehicle," in *Proceedings of the 2002 American Control Conference (IEEE Cat. No. CH37301)*, pp. 3129–3134, Anchorage, AK, USA, May 2002.
- [46] J. K. Ahn, K. H. Jung, D. H. Kim, H. B. Jin, H. S. Kim, and S. H. Hwang, "Analysis of a regenerative braking system for hybrid electric vehicles using an electro-mechanical brake," *International Journal of Automotive Technology*, vol. 10, no. 2, pp. 229–234, 2009.
- [47] Y. Gao and M. Ehsani, "Design and control methodology of plug-in hybrid electric vehicles," *IEEE Transactions on Industrial Electronics*, vol. 57, no. 2, pp. 633–640, 2009.
- [48] X. Huang and J. Wang, "Model predictive regenerative braking control for lightweight electric vehicles with in-wheel motors," *Proceedings of the Institution of Mechanical Engineers - Part D: Journal of Automobile Engineering*, vol. 226, no. 9, pp. 1220–1232, 2012.
- [49] H. A. Borhan and A. Vahidi, "Model predictive control of a power-split hybrid electric vehicle with combined battery and ultracapacitor energy storage," in *Proceedings of the 2010 American Control Conference*, pp. 5031–5036, Baltimore, MA, USA, July 2010.
- [50] R. Ghorbani, E. Bibeau, P. Zanetel, and A. Karlis, "Modeling and simulation of a series parallel hybrid electric vehicle using REVS," in *Proceedings of the 2007 American Control Conference*, pp. 4413–4418, New York, NY, USA, July 2007.
- [51] J. Jinming Liu and H. Huei Peng, "Modeling and control of a power-split hybrid vehicle," *IEEE Transactions on Control Systems Technology*, vol. 16, no. 6, pp. 1242–1251, 2008.
- [52] Y. Song and B. Wang, "Analysis and experimental verification of a fault-tolerant HEV powertrain," *IEEE Transactions on Power Electronics*, vol. 28, no. 12, pp. 5854–5864, 2013.
- [53] L. Parsa and H. A. Toliyat, "Fault-tolerant interior-permanent-magnet machines for hybrid electric vehicle applications," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 4, pp. 1546–1552, 2007.
- [54] B. Akin, S. B. Ozturk, and H. A. Toliyat, "On-board fault diagnosis of HEV induction motor drive at start-up and during idle mode," in *Proceedings of the 2007 IEEE Vehicle Power and Propulsion Conference*, pp. 140–147, Arlington, TX, USA, September 2007.
- [55] K. Rothenhagen and F. W. Fuchs, "Current sensor fault detection, isolation, and reconfiguration for doubly fed induction generators," *IEEE Transactions on Industrial Electronics*, vol. 56, no. 10, pp. 4239–4245, 2009.
- [56] R. Jayabalan and B. Fahimi, "Monitoring and fault diagnosis of multiconverter systems in hybrid electric vehicles," *IEEE Transactions on Vehicular Technology*, vol. 55, no. 5, pp. 1475–1484, 2006.
- [57] R. Merzouki, M. A. Djeziri, and B. Ould-Bouamama, "Intelligent monitoring of electric vehicle," in *Proceedings of the 2009 IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, pp. 797–804, Singapore, July 2009.
- [58] S. Grammatico, A. Balluchi, and E. Cosoli, "A series-parallel hybrid electric powertrain for industrial vehicles," in *Proceedings of the 2010 IEEE vehicle power and propulsion conference*, pp. 1–6, Lille, France, September 2010.
- [59] M. Ehsani, Y. Gao, S. E. Gay, and A. Emadi, *Modern Electric, Hybrid Electric, and Fuel Cell Vehicles*, CRC Press, Boca Raton, FL, USA, 2018.
- [60] C. Sankavaram, B. Pattipati, K. R. Pattipati, Y. Yilu Zhang, and M. Howell, "Fault diagnosis in hybrid electric vehicle regenerative braking system," *IEEE Access*, vol. 2, pp. 1225–1239, 2014.
- [61] Y. Hu, S. Zeng, L.-A. Carlos, K. Ullah, and Y. Yang, "Social network group decision-making method based on Q-rung Orthopair fuzzy set and its application in the evaluation of online teaching quality," *Axioms*, vol. 10, no. 3, p. 168, 2021.



- [62] H. Garg, "A new possibility degree measure for interval-valued q-rung orthopair fuzzy sets in decision-making," *International Journal of Intelligent Systems*, vol. 36, no. 1, pp. 526–557, 2021.
- [63] T. Mahmood and U. Rehman, "A novel approach towards bipolar complex fuzzy sets and their applications in generalized similarity measures," *International Journal of Intelligent Systems*, vol. 37, no. 1, pp. 535–567, 2022.
- [64] Z. Mu, S. Zeng, and P. Wang, "Novel approach to multi-attribute group decision-making based on interval-valued Pythagorean fuzzy power Maclaurin symmetric mean operator," *Computers & Industrial Engineering*, vol. 155, Article ID 107049, 2021.
- [65] Z. Ali, T. Mahmood, T. Mahmood, K. Ullah, and Q. Khan, "Einstein geometric aggregation operators using a novel complex interval-valued pythagorean fuzzy setting with application in green supplier chain management," *Reports in Mechanical Engineering*, vol. 2, no. 1, pp. 105–134, 2021.
- [66] D. Bozanic, D. Tešić, D. Tešić, D. Marinkovic, and A. Milić, "Modeling of neuro-fuzzy system as a support in decision-making processes," *Reports in Mechanical Engineering*, vol. 2, no. 1, pp. 222–234, 2021.
- [67] A. R. Mishra, P. Liu, and P. Rani, "COPRAS method based on interval-valued hesitant Fermatean fuzzy sets and its application in selecting desalination technology," *Applied Soft Computing*, vol. 119, Article ID 108570, 2022.
- [68] J. M. Mendel, "On computing the similarity of trapezoidal fuzzy sets using an automated area method," *Information Sciences*, vol. 589, pp. 716–737, 2022.

## Research Article

# Cross-Project Defect Prediction Based on Two-Phase Feature Importance Amplification

Ying Xing <sup>1</sup>, Wanting Lin <sup>1</sup>, Xueyan Lin <sup>1</sup>, Bin Yang <sup>2</sup>, and Zhou Tan <sup>1</sup>

<sup>1</sup>School of Artificial Intelligence, Beijing University of Posts and Telecommunications, 100876 Beijing, China

<sup>2</sup>Du Xiaoman (Beijing) Science Technology Co., Ltd., 100000 Beijing, China

Correspondence should be addressed to Wanting Lin; [wantinglin@bupt.edu.cn](mailto:wantinglin@bupt.edu.cn)

Received 24 January 2022; Accepted 4 March 2022; Published 18 April 2022

Academic Editor: Daqing Gong

Copyright © 2022 Ying Xing et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

As the typical application of computational intelligence in software engineering, cross-project defect prediction (CPDP) uses labeled data from other projects (source projects) for building models to predict the defects in the current projects (target projects), helping testers quickly locate the defective modules. But class imbalance and different data distribution among projects make CPDP a challenging topic. To address the above two problems, we propose a two-phase feature importance amplification (TFIA) CPDP model in this paper which can solve these two problems from domain adaptation phase and classification phase. In the domain adaptation phase, the differences in data distribution among projects are reduced by filtering both source and target projects, and the correlation-based feature selection with greedy best-first search amplifies the importance of features with strong feature-class correlation. In the classification phase, Random Forest works as the classifier to further amplify the importance of highly correlated features and establish a model which is sensitive to highly correlated features. We conducted both ablation experiments and comparison experiments on the widely used AEEEM database. Experimental results show that TFIA can yield significant improvement on CPDP. And the performance of TFIA CPDP model in all experiments is stable and efficient, which lays a solid foundation for its further application in practical engineering.

## 1. Introduction

Software defects may cause unexpected disasters in the application scenarios, threatening the security of software and even people's lives [1]. During the whole life cycle of software, defects that are found in the later period will cost more than those found in the earlier period to be repaired. Testers should find out the defects in time, helping developers fix them as soon as possible [2]. But focusing on all modules of the software will cost a large amount of time and manpower, which goes against our original intention and makes preidentification of potentially defective modules an urgent issue [3, 4].

Software defect prediction (SDP) is a method that can assist testers in quickly identifying potentially defective modules [5] so as to reduce the time spent by testers on troubleshooting and save testing costs as well [4]. The typical SDP usually uses historical data of software to build

prediction models for defect pattern recognition in new releases of the modules from the projects [3, 6, 7]. As the brand-new projects lack historical data, especially the historical data with defective or clean labels, SDP is unable to work well on them. To solve this problem, researchers propose a method called cross-project software defect prediction (CPDP). CPDP builds prediction models with historical data from other projects (source projects) and predicts the current projects (target projects) [8, 9].

Although CPDP focuses on the problem of lack of historical data, there are still two major problems to be solved. Due to a variety of factors such as project functionality and developers' habits, data distribution varies greatly from project to project. Differences in data distribution between projects are evident even when the same metrics are used to evaluate projects [10], which makes CPDP models built on source projects ineffective on the target project [11]. Therefore, how to reduce the differences



in data distribution between source and target projects is important to the performance of CPDP models [12, 13]. On the other hand, the number of the modules with defects is often much smaller than that of the modules without defects, which is called class imbalance problem [14]. The class imbalance problem can also affect the performance of CPDP models, because they may have a preference for the majority when classifying [15, 16].

To address the above problems, in this paper we propose a two-phase feature importance amplification (TFIA) CPDP model. Specifically, TFIA divides CPDP into two phases. In the domain adaptation phase, TFIA reduces differences in data distribution among source and target projects. After adding a certain proportion of samples from target projects to the source projects, the correlation-based feature selection method with a greedy best-first search strategy is proposed to amplify the importance of features with high feature-class correlation. The subset made up of the final selected items is used to filter the source and target projects for the purpose of removing redundant features as well as interfering items and reducing the dimensionality of the data. In the classification phase, resampling is conducted on the source project's data to make the numbers of clean and defective samples balanced. And an integrated learning method, Random Forest, is adopted to build the prediction model. The reason for adopting Random Forest is that integrated learning methods have shown their advantages in solving the class imbalanced problem [17]. In our case, the selected Random Forest uses a random sampling method; thus the trained model has small variance and high generalization ability, which can improve and effectively solve the class imbalance problem. Random Forest evaluates the relationship between features and classes during the training process [18], resulting in a ranking of feature-class correlation [19], further amplifying the importance of features with strong feature-class correlation for defect recognition.

In summary, our contributions lie in twofold.

- (1) We propose a two-phase feature importance amplification CPDP model (TFIA). TFIA reduces the effect of data distribution and class imbalance problems in domain adaptation phase and classification phase.
- (2) We conducted experiments on the AEEEM database to verify the validity of the method. In addition, we performed a number of ablation experiments to carefully analyse and dissect the detailed components of our method. And we also conducted comparison experiments to compare TFIA with other methods.

## 2. Related Work

CPDP usually includes two phases, called domain adaptation and classification. Domain adaptation is mainly addressed by data processing methods, such as data migration and feature selection. Classification is mainly handled by machine learning classifiers [20]. The researches on domain adaptation address the difference in data

distribution between the source and target projects, while the researches on classification methods improve the prediction performance from the perspective of algorithms. And also some researchers consider hybrid methods.

*2.1. Domain Adaptation Phase.* Turhan et al. [21] proposed Burak filter based on k-means clustering method. Burak filter calculated the Euclidean distance between all samples in the source project datasets. And for each sample in the target project datasets, they selected the  $k$  samples with the smallest Euclidean distance as training samples. Finally, the Naive Bayes method was used to predict the processed dataset. Burak filter improves the performance of CPDP model by filtering the samples of the source project datasets based on the target project datasets. Peters et al. [22] proposed Peters filter based on the source project datasets. For the samples in the source project datasets, the sample which was the closest to it in the target project datasets was selected by comparing the Euclidean distance. Then they labeled that sample. The sample of source project datasets with the closest distance to the labeled sample was selected as the training data in the prediction process. Pan et al. [23] proposed the TCA, which mapped the features of the source and target project to the latent space that makes them most similar, reducing the effect of differences in data distribution. Finally, Logistic Regression was used for prediction. He et al. [24] simplified the training set by TDSelector method and then classified it by Logistic Regression. Sun et al. [25] proposed a near-some source project selection by collaborative filtering (CFPS) method to filter source items, which has good results using SMO and Random Forest as classifiers. Alsawalqah et al. [26] proposed a SMOTE-ensemble method to optimize for class imbalance problems at the data level and algorithm level.

*2.2. Hybrid Method.* Yuan et al. [27] used TrAdaBoost to determine weights for samples based on Burak filter and used weighted support vector machines to build the model to improve the CPDP model. Chao et al. [28] proposed a two-phase CPDP method called TPTL, using a source project estimator to select source projects with similar data distribution as the target project and using two improved TCA + to construct models for prediction. Cong [29] proposed a DA-KTSVMO method using kernel twin support vector machines to improve the data distribution and using a quantum particle swarm optimization algorithm to optimize the method for prediction. Zhang et al. investigated seven composite algorithms; they believed that composite algorithms can improve the performance of CPDP models [30]. The method proposed by Chen et al. combines the data gravitation method and TrAdaBoost to reduce the effects of class imbalance in the source project [31]. Xu et al. proposed a multisource TrAdaBoost Algorithm. The proposed method uses semisupervised high-density-based clustering and a small amount of labeled target item data to obtain a large amount of labeled source item data. When using TrAdaBoost for integration, only the base classifier trained from the source item data most relevant to the target item is

selected. In the process of training, this method can ensure that the knowledge transferred is most relevant to the target project but ignores the influence of other source project data on the target project defect prediction model [32].

### 3. Methodology

In this section, we first introduce the framework of our proposed method TFIA. Then we show the details of important steps in our proposed method.

**3.1. Framework of TFIA.** Figure 1 is the flowchart of TFIA. TFIA adds a certain proportion of labeled samples from the target project datasets to the source project datasets, uses a correlation-based feature selection method, and searches for a subset of features using a heuristic greedy best-first search strategy. Then TFIA filters the feature of source and target projects, samples the source projects datasets with resampling, uses Random Forest on the source project datasets for training, and finally predicts the target project datasets.

TFIA reduces the difference in data distribution between source and target projects from the feature perspective and sample perspective, respectively. TFIA enhances the importance of features with strong class-feature correlation in model training. A resample approach is used to deal with class imbalance problems and reduce the impact generated by class imbalance problem. TFIA uses Random Forest as a classifier to train models with small variance and strong generalization ability to improve the accuracy of CPDP model for identifying defective modules.

In TFIA, we use Relief to evaluate the correlation of feature-feature and feature-class in TFIA. Relief[33] is a feature weighting algorithm that is sensitive to feature interactions. Relief of feature  $f$  can be formulated as

$$\text{Relief}_f = \frac{\text{Gini}' \times \sum_{f_v \in f} p(f_v)^2}{\left(1 - \sum_{c \in C} p(c)^2\right) \sum_{c \in C} p(c)^2}. \quad (1)$$

$p$  is the probability,  $C$  is the class variable,  $C = \{\text{defective}, \text{clean}\}$ ,  $f_v$  is the certain value of the sample's feature  $f$ , and  $\text{Gini}'$  is another attribute quality measure that can be calculated by

$$\begin{aligned} \text{Gini}' = & \left[ \sum_{c \in C} p(c)(1 - p(c)) \right] \\ & - \sum_{f_v \in f} \left( \frac{p(f_v)^2}{\sum_{f_v \in f} p(f_v)^2} p(f_v)^2 \sum_{c \in C} p(c f_v)(1 - p(c f_v)) \right). \end{aligned} \quad (2)$$

We use Relief to calculate correlation of feature-feature and feature-class in Section 3.2 and Section 3.3.

**3.2. Domain Adaptation Phase.** In order to amplify the role of features with strong feature-class correlation in the classifier and to deemphasize features with strong feature-feature correlation, we decide to select the features. In the feature selection phase, we use the filtering method called correlation-based feature selection (CFS) [34]. In this paper,

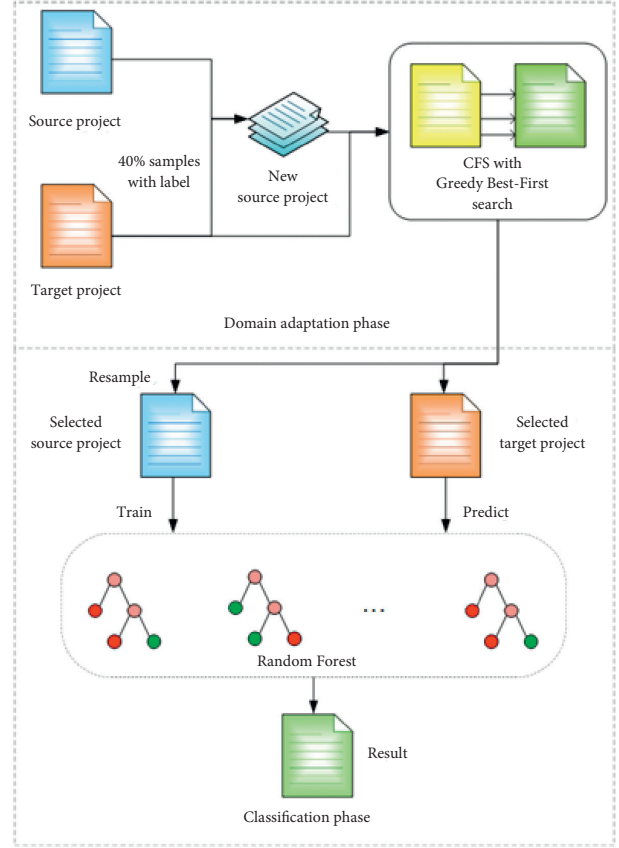


FIGURE 1: Flowchart of TFIA.

the heuristic search strategy used by CFS is greedy best-first search [35]. CFS evaluates the value of a subset of features by considering the individual predictive ability of each feature and the degree of redundancy between them [36].

As shown in Figure 2, CFS first computes the feature-class and feature-feature correlation matrices from the source project dataset and then searches the feature subset space using greedy best-first search.

To prevent the best-first search from exploring the entire feature subset search space, we follow the setting of [34] to impose a termination criterion. The search will terminate if five consecutive fully expanded subsets show no improvement over the current best subset. CFS filters features by a feature subset evaluation function, intending to find a subset of features that meet the conditions of low feature-feature correlation and strong feature-class correlation, thus sifting out redundant features [37]. The feature subset evaluation function is valued as the merit which can be calculated by

$$M_S = \frac{k \overline{r_{fc}}}{\sqrt{k + k(k-1) \overline{r_{ff}}}}, \quad (3)$$

$M_S$  is the value of the merit of feature subset with  $k$  selected features.  $\overline{r_{fc}}$  is the average feature-class correlation and  $\overline{r_{ff}}$  is the average feature-feature correlation.  $r$  is Relief in Section 3.1.

The algorithm is described as follows.

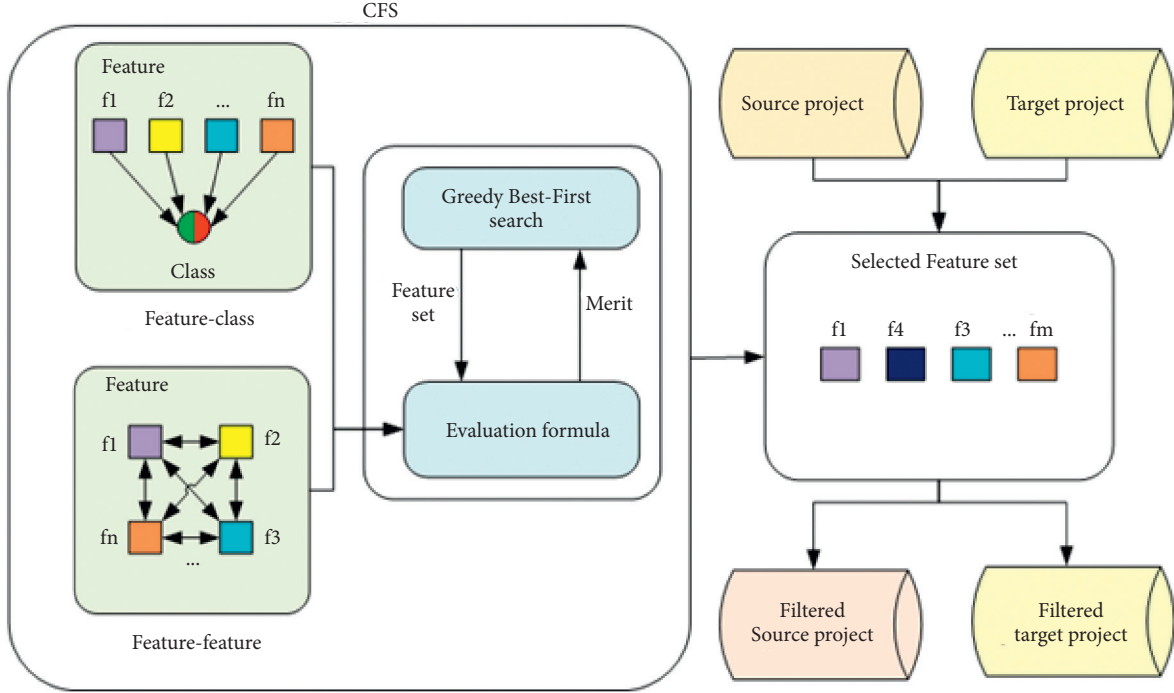


FIGURE 2: Flowchart of CFS.

3.3. *Classification Phase.* Random Forest is an integrated learning pattern recognition method [19, 38]. It has been demonstrated that Random Forest has good performance in CPDP [27, 39, 40] due to its high tolerance to outliers and noise. Random Forest is also less prone to fitting characteristics [18].

Figure 3 shows the process of Random Forest in this paper. Random Forest uses a bootstrap method to randomly scrape multiple samples from the original samples, modeling a Decision Tree for each sample set. Then we give a comprehensive conclusion from the results of all decision trees. The voting decision process of Random Forest is shown by

$$H(x) = \text{majority vote} \left\{ \max_C \sum_{i=1}^k I(h_i(x) = C) \right\}, \quad (4)$$

$H(x)$  denotes the combined classification model,  $h_i$  denotes the single subject Decision Tree,  $C$  is the set of class labels, and  $I(\cdot)$  is the indicative function.

There is a put-back from the source project dataset  $D$  processed in the domain adaptation phase to obtain a randomly selected subset  $D_{T_i}$  as the training set. And the sample size of the training dataset is the same as the original data.

This sampling method of Random Forest ensures the variability of the training set. For the sake of easy explanation, we assume that there are  $n$  samples in dataset  $D$ . The probability of each sample being picked is  $1/n$ . Repeat it for  $n$  times, so the probability of each sample in  $D$  not being picked is  $(1 - 1/n)^n$ . As  $n$  tends to infinity,  $\lim_{n \rightarrow \infty} (1 - 1/n)^n = 0.368$ . We can assume that 37% of the samples in  $D$  will not appear, which guarantees the variability of the training set.  $\lambda$  features are randomly selected

from the source project dataset to construct a Decision Tree, and each node is based on (2).

In a single Decision Tree, the *Gini'* metric is calculated for each attribute and a variable with the minimum *Gini'* metric is selected to split the current node. The Decision Tree is constructed by recursion until the stop criterion is reached.

The algorithm is described as follows.

## 4. Experimental Design

For our method TFIA, in this section, we raise four research questions and set up experiments to discuss and analyse each question.

4.1. *Research Question.* RQ1: Does the feature-level filtering approach proposed in this paper have any impact on the performance of the model?

In TFIA, we use filtering to reduce the data distribution difference in source and target projects datasets by CFS. Therefore we intend to analyse the role of filtering methods in the whole CPDP model.

RQ2: In the classification phase, does the choice of classifier affect the overall performance?

In TFIA, feature-class correlation is also amplified in the classification phase as we use Random Forest as the classifier. Classifiers behave diversely on different types of data, so we want to investigate how other classifiers would perform on our model, such as the linear classifiers Logistic Regression and Support Vector Classification, the Bayesian formula-based classifier Naive Bayes, and the tree-based classifier Decision Tree.

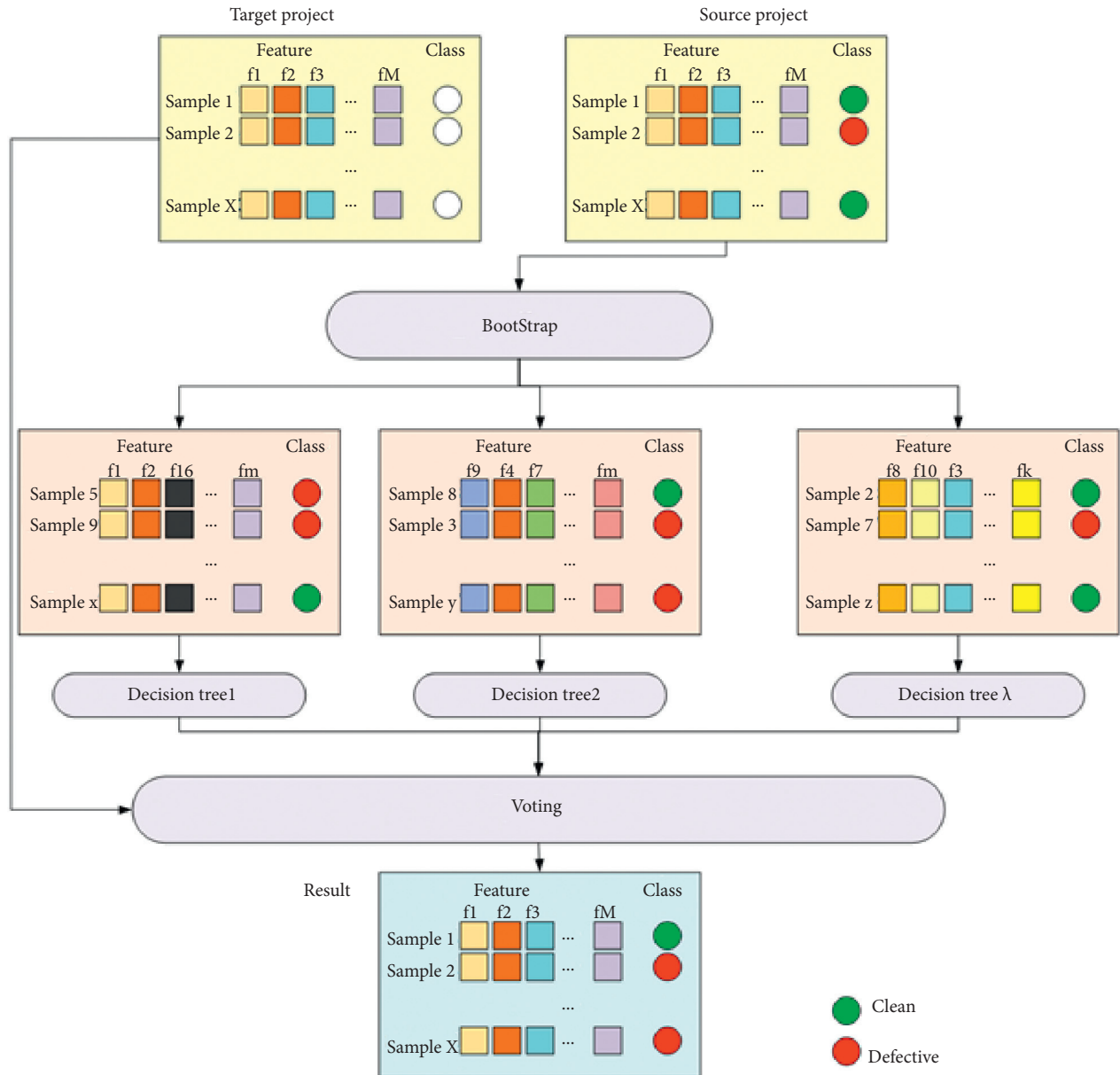


FIGURE 3: Flowchart of Random Forest.

RQ3: When features are transferred at the sample level, does the proportion of transfer affect the results?

In the domain adaptation phase of TFIA, a certain percentage of samples of the labeled target items are required. We want to analyse how different proportions of samples of labeled target items in the project affect the overall prediction performance.

RQ4: Compared with the classical method and the latest research, does the method proposed in this paper improve CPDP model's performance?

Various methods have been proposed by researchers for the CPDP problem, such as the classical methods TCA [23], Peters filtering method [22], Burak filtering method [21], and the newer method ALTRA [27]. We seek to analyse how the performance obtained by TFIA differs from these methods.

4.2. *Dataset.* In this paper, we use the AEEEM database, a widely used dataset in the field of software defect prediction research. The AEEEM database was collected and compiled by D'Ambrosio et al. [41]. Table 1 shows the details of the AEEEM database, including the projects names, the projects types, numbers of modules, numbers of defective modules, and the ratio of defective modules.

4.3. *Experimental Environment.* All of our codes are written in *Python 3.7*. The GPU used for the experiments is NVIDIA TITAN V and the CPU is Intel I9-9920X. Classifications are realized by WEKA [42] with default parameters.

4.4. *Performance Measure.* The confusion matrix is used to store the correct and incorrect decisions made by the



(i) Input:  $F$ , Feature set with  $n$  features and a class;  $F = \{f_1, f_2, \dots, f_n, C\}$   
(ii) Output:  $S$ , Subset of selected features,  $S \in F$   
1  $S = \{\emptyset\}$ ,  $F_1 = \{f_1\}$ ,  $F_2 = \{f_2\}$ , ...,  $F_n = \{f_n\}$   
2 compute the merit of  $F_1 = \{f_1\}$ ,  $F_2 = \{f_2\}$ , ...,  $F_n = \{f_n\}$  by (3)  
3 put the feature with the biggest merit of  $F$  into  $S$ , and remove it from  $F$   
4 compute merit of  $S$  ( $M_S$ )  
5 put the new feature with the biggest merit of  $F$  into  $S$ , compute  $M_S$ ,  $\text{merit}_{\text{new}} = M_S$   
6 if  $\text{merit}_{\text{old}} \geq \text{merit}_{\text{new}}$  then  
7 remove the new feature from  $S$   
8 else  
9 Update  $S$ , remove the new feature from  $F$ ,  $\text{merit}_{\text{old}} = \text{merit}_{\text{new}}$   
10 end if  
11 repeat step 5 to step 7 until  $F = \{\emptyset\}$  or  $\text{merit}_{\text{old}}$  has no change by a certain number of times  
12 return  $S$

ALGORITHM 1: CFS.

prediction model. The purpose of our defect prediction is to test out defective data. In this paper, we believe that a sample itself is defective, and the classifier also considers it defective, called True Positive (TP). The classification is False Negative (FN) if the classifier thinks it is clean. Similarly, if a sample does not have defects, the classifier classifies it as defective, called False Positive (FP); the classifier classifies it as free of defects, called True Negative (TN).

**4.5. F1-Measure.** F1-measure, also known as F-score, is a weighted summed average of Precision and Recall. Precision is the ratio of correctly predicted defective modules to all modules predicted to be defective, calculated as shown by

$$\text{Precision} = \frac{TP}{TP + FP}. \quad (5)$$

Recall is the ratio of correctly predicted defective modules to all truly defective modules, calculated as shown in

$$\text{Recall} = \frac{TP}{TP + FN}. \quad (6)$$

F1-measure is often used to evaluate the performance of a classification model. F1-measure can be calculated by

$$F1 - \text{measure} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} = \frac{2TP}{2TP + FP + FN}. \quad (7)$$

**4.6. Area under the ROC Curve.** Area under the ROC curve (AUC) is used to evaluate the degree of discrimination obtained by the model. The value of AUC ranges in  $[0, 1]$ . AUC for random prediction is 0.5. The advantage of AUC is that AUC is insensitive to decision thresholds such as precision and recall. The higher the AUC, the better the prediction.

**4.7. Matthews Correlation Coefficient.** Matthews correlation coefficient (MCC) is used in machine learning as a measure

of binary (2-category) quality of classification, which is introduced from biochemistry by Brian W. Matthews in 1975 [43].

MCC takes into account true and false, positives and negatives, and is generally regarded as a balanced measure that can be used even if the classes are of very different sizes [44]. MCC is essentially a correlation coefficient between the observed and predicted binary classification. It returns a value between -1 and +1.  $MCC = +1$  indicates a perfect prediction,  $MCC = 0$  indicates that the model is not better than a random prediction, and  $MCC = -1$  indicates a complete inconsistency between prediction and observation.

The formula for the MCC is shown by

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}. \quad (8)$$

## 5. Results and Discussion

In this section, we present the experimental results and give the answers to questions from Section 3.

**5.1. Answer for RQ1.** To answer RQ1, we designed two groups of experiments on the AEEEM database, where we formed a total of 20 source-target project pairs from the five projects, including EQ, JDT, ML, LC, and PDE. The first group of experiments performed no filtering on these 20 source-target project pairs. And the second group of experiments were filtered by CFS based on greedy best-first search strategy proposed in this paper. Both sets of experiments used Random Forest as the classifier. Both groups of experiments randomly added 40% of the target project data to the source project dataset and source project datasets were dealt with through resample.

Table 2 shows the results of the evaluation metrics obtained on 20 source-target project pairs without (Model 1) and with (Model 2) the filtering method proposed in this paper.

(i) Input:  $D_S$ , source project dataset;  $\text{sample}_T$ , sample from source project dataset  
 Output:  $C(\text{sample}_T)$  = classification of  $\text{sample}_T$

- (1) for  $\{i = 1, 2, \dots, \lambda\}$  do
- (2) use Bootstrap on  $D_S$  to get the training dataset  $D_{T_i}$
- (3) use  $D_{T_i}$  to generate a tree  $h_i$  without pruning
- (4) randomly select  $d$  features from  $D_S$ 's features
- (5) calculate the  $Gini'$  metric based on (2) at each node to select the optimal features based on the selected  $d$  features
- (6) Splitting until the tree grows to its maximum
- (7) end for
- (8) return  $C(\text{sample}_T) = \text{majority vote}\{h_i(\text{sample}_T)\}_{i=1}^{\lambda_{\text{tree}}}$

ALGORITHM 2: Random Forest.

TABLE 1: Details of AEEEM database.

Projects	Types of projects	Modules	Defective modules	Defective ratio (%)
EQ	OSGI framework	325	129	39.692
JDT	Development	997	206	20.662
LC	Text search engine library	399	64	16.040
ML	Task management	1862	245	13.158
PDE	Development	1492	209	14.008

TABLE 2: F1-measure, AUC, and MCC of model with filtering (Model 1) and model without filtering (Model 2).

	F1-measure		AUC		MCC	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
EQ-JDT	0.415	<b>0.790</b>	0.632	<b>0.873</b>	0.232	<b>0.735</b>
EQ-LC	0.300	<b>0.761</b>	0.726	<b>0.903</b>	0.265	<b>0.738</b>
EQ-ML	0.242	<b>0.742</b>	0.545	<b>0.875</b>	0.061	<b>0.702</b>
EQ-PDE	0.291	<b>0.776</b>	0.600	<b>0.864</b>	0.146	<b>0.740</b>
JDT-EQ	0.276	<b>0.833</b>	0.576	<b>0.862</b>	0.291	<b>0.719</b>
JDT-LC	0.322	<b>0.752</b>	0.602	<b>0.875</b>	0.330	<b>0.726</b>
JDT-ML	0.283	<b>0.745</b>	0.584	<b>0.883</b>	0.229	<b>0.707</b>
JDT-PDE	0.247	<b>0.761</b>	0.569	<b>0.866</b>	0.233	<b>0.721</b>
LC-EQ	0.196	<b>0.831</b>	0.554	<b>0.862</b>	0.261	<b>0.710</b>
LC-JDT	0.592	<b>0.816</b>	0.727	<b>0.911</b>	0.516	<b>0.768</b>
LC-ML	0.347	<b>0.765</b>	0.626	<b>0.855</b>	0.243	<b>0.732</b>
LC-PDE	0.256	<b>0.787</b>	0.572	<b>0.866</b>	0.234	<b>0.755</b>
ML-EQ	0.183	<b>0.855</b>	0.550	<b>0.886</b>	0.251	<b>0.755</b>
ML-JDT	0.396	<b>0.787</b>	0.623	<b>0.883</b>	0.397	<b>0.730</b>
ML-LC	0.219	<b>0.824</b>	0.562	<b>0.925</b>	0.316	<b>0.806</b>
ML-PDE	0.217	<b>0.763</b>	0.559	<b>0.865</b>	0.229	<b>0.724</b>
PDE-EQ	0.238	<b>0.873</b>	0.560	<b>0.897</b>	0.232	<b>0.785</b>
PDE-JDT	0.477	<b>0.806</b>	0.666	<b>0.881</b>	0.366	<b>0.754</b>
PDE-LC	0.372	<b>0.784</b>	0.620	<b>0.873</b>	0.397	<b>0.763</b>
PDE-ML	0.292	<b>0.754</b>	0.594	<b>0.853</b>	0.178	<b>0.718</b>
<b>Average</b>	0.308	<b>0.790</b>	0.602	<b>0.878</b>	0.270	<b>0.739</b>

The values in bold are results with the best performance of each instance.

From Table 2, it can be observed that the overall effect of the prediction models trained from the data processed by the filtering method is significantly improved (the data in bold), with an average F1-measure improvement of about 156.54%, AUC improvement of about 45.74%, and MCC improvement of about 173.46%. Numerically, in these 20 source-target project pairs, the models with filtering all work better than those without filtering.

Figure 4 is the boxplot of F1-measure, AUC, and MCC of model with filtering (Model 1) and model without filtering

(Model 2) based on data from Table 2. As can be seen from Figure 4, the prediction models trained from the data processed by the filtering have higher numerical intervals in the overall distribution compared to the data without filtering. In terms of box size and endline length, the overall performance of the model after processing with the filtering also becomes less volatile and the performance is more stable. Therefore, we can conclude that the filtering method proposed in this paper is effective in improving the performance of the CPDP model on these datasets.

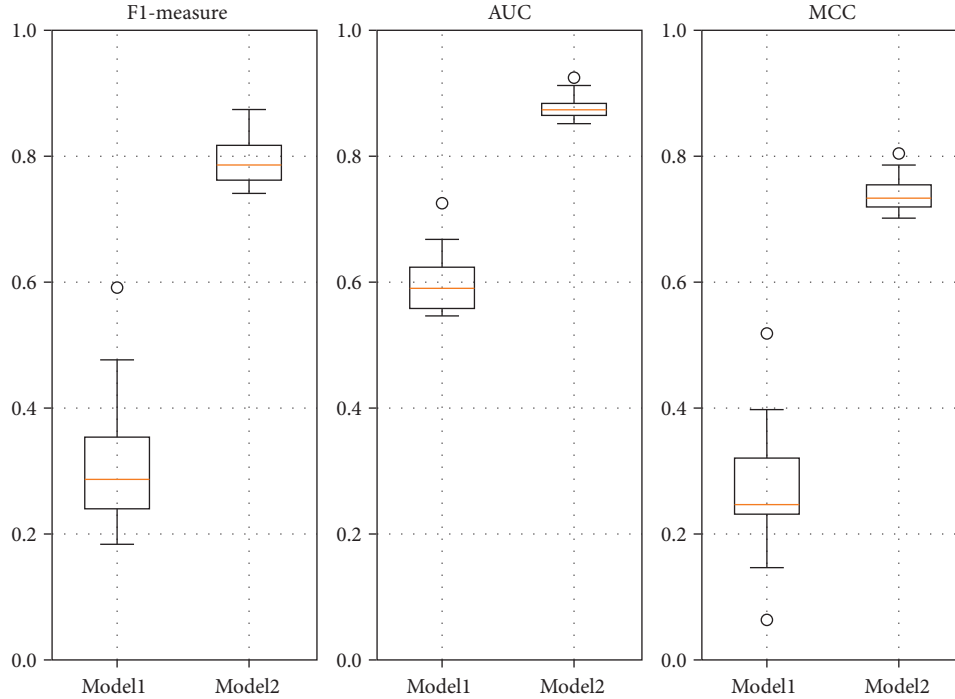


FIGURE 4: Boxplot of F1-measure, AUC, and MCC of model with filtering (Model 1) and model without filtering (Model 2).

5.2. *Answer for RQ2.* To answer RQ2, 20 source-target projects from the AEEEM database were trained and validated using Naive Bayes (NB), Logistic Regression (LR), Decision Tree (DT), Support Vector Classification (SVC), and Random Forest (RF) as classifiers. Respectively, all the datasets were processed by the filtering method proposed in this paper. Every set of experiments randomly added 40% sample from the target project datasets to the source project datasets and the source project datasets were dealt with through resample.

Table 3 shows the F1-measure, AUC, and MCC of models with different classifications, including Naive Bayes (NB), Logistic Regression (LR), Decision Tree (DT), Support Vector Classification (SVC), and Random Forest (RF). The data in bold show the classification with the best performance in each set of experiments. The value of F1-measure shows that, on the data processed by our filtering method in this paper, RF has the best results, Decision Tree is the second best, and LR, NB, and SVC are more similar and differ from RF and DT. The situation is the same in AUC and MCC. This phenomenon could be caused by the idea that, in the first phase, our filtering method amplifies the proportion of features that have a strong correlation with the class. So it has better performance on DT and Decision Tree-based integrated method RF. The other three classifiers are relatively insensitive to the dataset of amplified features, so the results are worse than DT and RF.

Figure 5 is the boxplots of F1-measure, AUC, and MCC of models with different classifications, including Naive Bayes (NB), Logistic Regression (LR), Decision Tree (DT), Support Vector Classification (SVC), and Random Forest (RF). From Figure 5, it can be seen that the prediction models constructed with Random Forest

and Decision Tree as classifiers achieve higher performance metrics and less overall volatility. This indicates that the choice of classifier has an impact on the performance of the prediction model. On the dataset processed by the filtering method in this paper, classifiers that are more sensitive to feature importance will have better results.

5.3. *Answer for RQ3.* To analyse the effect of the proportion of samples from target projects datasets added to the source projects datasets on the prediction results, we added different proportions of labeled target item samples to the source projects datasets in steps of 10% from 0% and trained with TFIA. Every source project dataset had been dealt with through resample.

Table 4 shows the average F1-measure, AUC, and MCC values of models added samples from target projects to the source projects at different proportions on 20 source-target project pairs.

Then we smoothed the values to plot the graphs as Figure 6, from which we can see that the transformation of the three measures starts to become very small at around 70%, indicating that the model performance reaches its best at around 70%. But for the consideration of the actual prediction environment, it is not feasible to manually label the target projects samples at 70%, so we only consider the growth rate of indicators in this paper. It can be seen that the AUC starts to level off around 40%, and the improvement rates of F1-measure and MCC start to become smaller, so we believe that adding 40% of the labeled samples of the target projects to the source projects will have better results in these 20 source-target projects pairs.



TABLE 3: F1-measure, AUC, and MCC of models with different classifications, including Naive Bayes (NB), Logistic Regression (LR), Decision Tree (DT), Support Vector Classification (SVC), and Random Forest (RF).

	F1-measure					AUC					MCC				
	LR	DT	NB	SVC	RF	LR	DT	NB	SVC	RF	LR	DT	NB	SVC	RF
EQ-JDT	0.594	0.729	0.560	0.571	<b>0.790</b>	0.758	0.852	0.713	0.727	<b>0.873</b>	0.479	0.655	0.467	0.465	<b>0.735</b>
EQ-LC	0.435	0.619	0.404	0.224	<b>0.761</b>	0.675	0.889	0.654	0.575	<b>0.903</b>	0.387	0.601	0.361	0.136	<b>0.738</b>
EQ-ML	0.375	0.672	0.333	0.376	<b>0.742</b>	0.647	0.866	0.611	0.637	<b>0.875</b>	0.271	0.627	0.253	0.288	<b>0.702</b>
EQ-PDE	0.379	0.660	0.361	0.386	<b>0.776</b>	0.664	0.851	0.621	0.667	<b>0.864</b>	0.259	0.607	0.283	0.269	<b>0.740</b>
JDT-EQ	0.595	0.829	0.492	0.459	<b>0.833</b>	0.680	0.857	0.644	0.632	<b>0.862</b>	0.379	<b>0.727</b>	0.358	0.347	0.719
JDT-LC	0.362	0.573	0.373	0.337	<b>0.752</b>	0.694	0.837	0.672	0.687	<b>0.875</b>	0.293	0.539	0.303	0.266	<b>0.726</b>
JDT-ML	0.376	0.637	0.361	0.418	<b>0.745</b>	0.644	0.854	0.632	0.679	<b>0.883</b>	0.276	0.589	0.265	0.320	<b>0.707</b>
JDT-PDE	0.373	0.651	0.356	0.381	<b>0.761</b>	0.668	0.841	0.619	0.679	<b>0.866</b>	0.253	0.595	0.279	0.265	<b>0.721</b>
LC-EQ	0.672	0.806	0.578	0.729	<b>0.831</b>	0.727	0.839	0.678	0.773	<b>0.862</b>	0.453	0.665	0.389	0.542	<b>0.710</b>
LC-JDT	0.575	0.721	0.585	0.525	<b>0.816</b>	0.755	0.860	0.743	0.730	<b>0.911</b>	0.451	0.648	0.473	0.384	<b>0.768</b>
LC-ML	0.393	0.655	0.350	0.406	<b>0.765</b>	0.647	0.843	0.620	0.659	<b>0.855</b>	0.307	0.604	0.272	0.314	<b>0.732</b>
LC-PDE	0.394	0.660	0.367	0.378	<b>0.787</b>	0.675	0.843	0.628	0.665	<b>0.866</b>	0.279	0.605	0.278	0.258	<b>0.755</b>
ML-EQ	0.696	0.826	0.560	0.632	<b>0.855</b>	0.730	0.857	0.674	0.687	<b>0.886</b>	0.455	0.703	0.397	0.369	<b>0.755</b>
ML-JDT	0.567	0.725	0.580	0.539	<b>0.787</b>	0.748	0.855	0.728	0.742	<b>0.883</b>	0.440	0.650	0.483	0.404	<b>0.730</b>
ML-LC	0.343	0.597	0.422	0.338	<b>0.824</b>	0.686	0.889	0.681	0.656	<b>0.925</b>	0.273	0.583	0.363	0.261	<b>0.806</b>
ML-PDE	0.401	0.653	0.381	0.395	<b>0.763</b>	0.684	0.837	0.635	0.680	<b>0.865</b>	0.288	0.597	0.295	0.281	<b>0.724</b>
PDE-EQ	0.640	0.791	0.497	0.525	<b>0.873</b>	0.704	0.826	0.645	0.638	<b>0.897</b>	0.410	0.643	0.357	0.299	<b>0.785</b>
PDE-JDT	0.585	0.701	0.602	0.521	<b>0.806</b>	0.746	0.834	0.752	0.733	<b>0.881</b>	0.470	0.619	0.495	0.382	<b>0.754</b>
PDE-LC	0.382	0.584	0.438	0.362	<b>0.784</b>	0.702	0.857	0.690	0.639	<b>0.873</b>	0.315	0.557	0.380	0.306	<b>0.763</b>
PDE-ML	0.385	0.629	0.312	0.406	<b>0.754</b>	0.646	0.831	0.600	0.666	<b>0.853</b>	0.291	0.573	0.232	0.308	<b>0.718</b>
<b>Average</b>	<b>0.476</b>	<b>0.686</b>	<b>0.446</b>	<b>0.445</b>	<b>0.790</b>	<b>0.694</b>	<b>0.851</b>	<b>0.662</b>	<b>0.678</b>	<b>0.878</b>	<b>0.351</b>	<b>0.619</b>	<b>0.349</b>	<b>0.323</b>	<b>0.739</b>

The data in bold shows the classification with the best performance in each set of experiments.

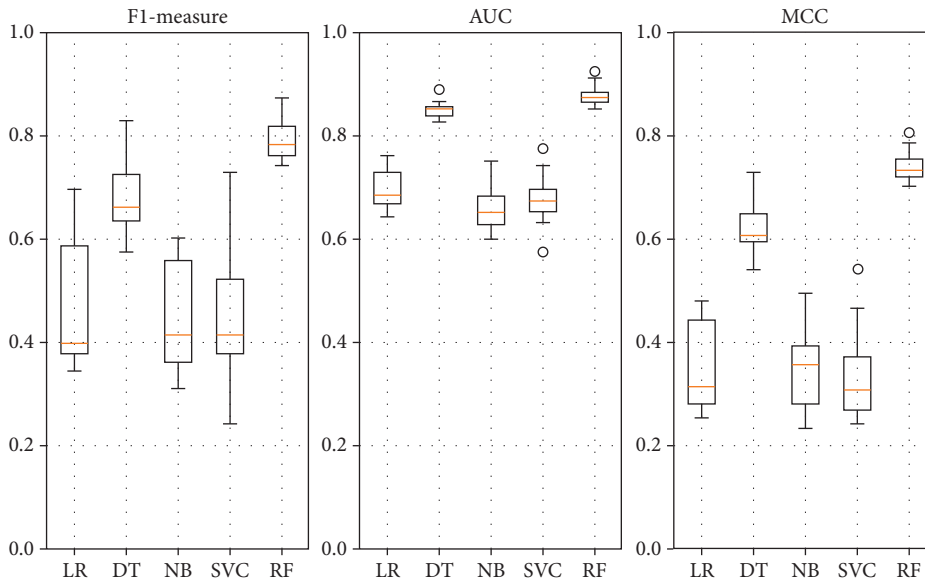


FIGURE 5: Boxplot of F1-measure, AUC, and MCC of models with different classifications, including Naive Bayes (NB), Logistic Regression (LR), Decision Tree (DT), Support Vector Classification (SVC), and Random Forest (RF).

TABLE 4: Average F1-measure, AUC, and MCC of TFIA using different proportions samples from target project datasets added to source project datasets.

Proportion (%)	Average F1-measure	Average AUC	Average MCC
0	0.352	0.609	0.221
10	0.458	0.668	0.350
20	0.575	0.736	0.475
30	0.683	0.804	0.607

TABLE 4: Continued.

Proportion (%)	Average F1-measure	Average AUC	Average MCC
40	0.790	0.878	0.739
50	0.866	0.928	0.832
60	0.921	0.963	0.902
70	0.953	0.981	0.942
80	0.969	0.990	0.962
90	0.980	0.995	0.976

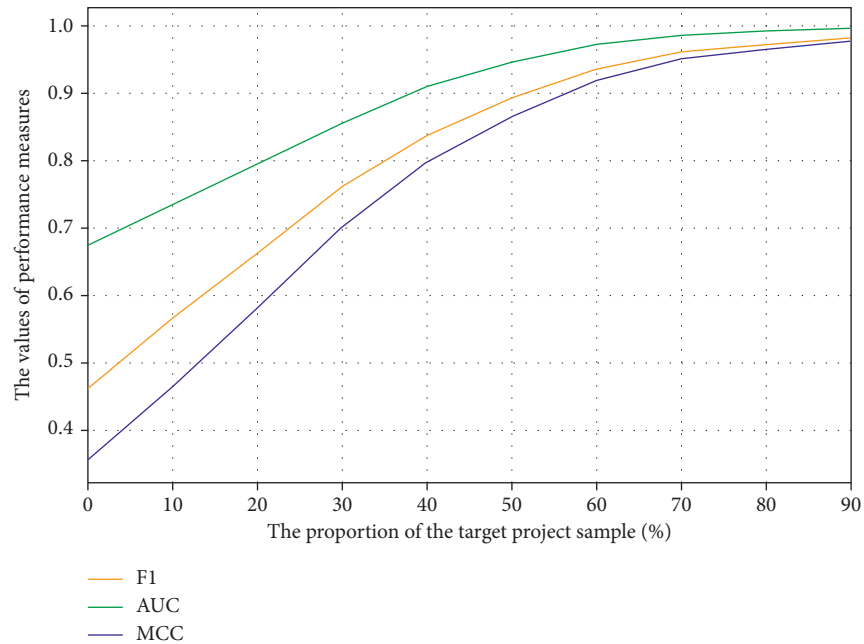


FIGURE 6: Average F1-measure, AUC, and MCC of TFIA using different proportions samples from target project datasets adding to source project datasets.

5.4. *Answer for RQ4.* To answer RQ4, we used TCA, Burak filter, Peters filter, and ALTRA on 20 source-target project pairs in the AEEEM database for prediction and compared with our proposed method, respectively. To ensure the consistency of the experimental conditions, the source items in the dataset were randomly added to 40% samples of the target project datasets, and the source project datasets were dealt with through resample.

Table 5 shows the F1-measure, AUC, and MCC of ALTRA (Model 1), TCA (Model 2), Peters filter (Model 3), Burak filter (Model 4), and TFIA (Model 5).

As can be seen from the data in Table 5, compared with other methods, the proposed method in this paper shows significant improvements in F1-measure, AUC, and MCC. Taken as an example, the value of F1-measure is improved by about 126.62% over TCA, 98.64% over Burak filter, 86.46% over Peters filter, and by 31.12% over ALTRA.

Figure 7 is the boxplot of F1-measure, AUC, MCC of ALTRA (Model 1), TCA (Model 2), Peters filter (Model 3), Burak filter (Model 4), and TFIA (Model 5). From Figure 7, we can see that the method in this paper has smaller boxes and shorter endline lengths relative to the other methods, indicating that TFIA performs more consistently on the

datasets. Although ALTRA performs better than TCA, Burak filter, and Peters filter on the average F1-measure, the fluctuation of ALTRA is larger. Therefore, we can conclude that the method proposed in this paper possesses good stability and prediction effect on the AEEEM database compared to these methods.

5.5. *Threats to Validity.* In this paper, the threats to validity are mainly divided into internal and external validity.

The internal validity comes from the setting of classifier parameters. In this paper, we set the parameters defaulted by the references as well as the tool of WEKA, which may lead to some differences in the classification results. To mitigate this difference, we use the same default parameters provided by WEKA for the classification phase of our proposed method.

The external validity mainly comes from the dataset used. We use five project datasets from the publicly available AEEEM database and combine 20 source-target project pairs for our experiments. If other data sources are selected, especially those from real-world engineering datasets, different experimental results may be obtained.

TABLE 5: FI-measure, AUC, and MCC of ALTRA (Model 1), TCA (Model 2), Peters filter (Model 3), Burak filter (Model 4), and TFIA (Model 5).

	FI-measure					AUC					MCC				
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 1	Model 2	Model 3	Model 4	Model 5	Model 1	Model 2	Model 3	Model 4	Model 5
EQ-JDT	0.448	0.419	0.574	0.552	0.873	0.266	0.638	0.729	0.707	0.735	-0.079	0.244	0.467	0.462	0.735
EQ-LC	0.449	0.220	0.404	0.384	0.903	0.286	0.613	0.719	0.658	0.738	-0.064	0.131	0.342	0.323	0.733
EQ-ML	0.304	0.320	0.376	0.353	0.875	0.253	0.656	0.646	0.623	0.702	-0.275	0.211	0.274	0.267	0.699
EQ-PDE	0.415	0.306	0.407	0.371	0.864	0.203	0.624	0.657	0.635	0.740	-0.233	0.180	0.309	0.269	0.740
JDT-EQ	0.526	0.727	0.598	0.459	0.862	0.388	0.760	0.694	0.632	0.719	-0.109	0.517	0.432	0.347	0.719
JDT-LC	0.704	0.235	0.405	0.377	0.875	0.271	0.647	0.705	0.665	0.726	0.062	0.172	0.340	0.309	0.725
JDT-ML	0.725	0.319	0.375	0.361	0.883	0.605	0.659	0.647	0.632	0.707	0.142	0.216	0.271	0.265	0.702
JDT-PDE	0.713	0.289	0.407	0.388	0.866	0.668	0.599	0.661	0.649	0.721	0.051	0.151	0.305	0.281	0.721
LC-EQ	0.465	0.673	0.588	0.566	0.862	0.297	0.708	0.683	0.674	0.710	-0.152	0.410	0.397	0.389	0.701
LC-JDT	0.868	0.379	0.567	0.580	0.911	0.441	0.576	0.725	0.734	0.768	0.091	0.161	0.458	0.472	0.762
LC-ML	0.862	0.322	0.378	0.348	0.855	0.341	0.665	0.647	0.619	0.732	-0.017	0.224	0.277	0.265	0.731
LC-PDE	0.792	0.298	0.397	0.394	0.866	0.605	0.612	0.657	0.652	0.755	0.078	0.164	0.290	0.289	0.754
ML-EQ	0.710	0.719	0.618	0.543	0.886	0.579	0.746	0.705	0.668	0.755	0.115	0.507	0.446	0.400	0.740
ML-JDT	0.751	0.376	0.568	0.567	0.883	0.309	0.572	0.724	0.721	0.730	-0.099	0.149	0.461	0.464	0.727
ML-LC	0.808	0.224	0.374	0.406	0.925	0.540	0.634	0.703	0.682	0.806	0.019	0.162	0.308	0.341	0.804
ML-PDE	0.806	0.294	0.408	0.390	0.865	0.287	0.606	0.664	0.649	0.724	-0.018	0.158	0.302	0.286	0.724
PDE-EQ	0.644	0.695	0.614	0.513	0.897	0.373	0.721	0.705	0.657	0.785	-0.174	0.450	0.453	0.390	0.784
PDE-JDT	0.800	0.384	0.571	0.582	0.881	0.388	0.586	0.721	0.731	0.754	0.077	0.169	0.476	0.481	0.754
PDE-LC	0.800	0.225	0.402	0.370	0.873	0.375	0.636	0.727	0.659	0.763	0.068	0.163	0.343	0.303	0.763
PDE-ML	0.800	0.325	0.384	0.342	0.853	0.491	0.669	0.654	0.618	0.718	0.068	0.230	0.281	0.254	0.718
Average	0.670	0.387	0.471	0.442	0.878	0.398	0.646	0.689	0.663	0.739	-0.022	0.238	0.362	0.343	0.737

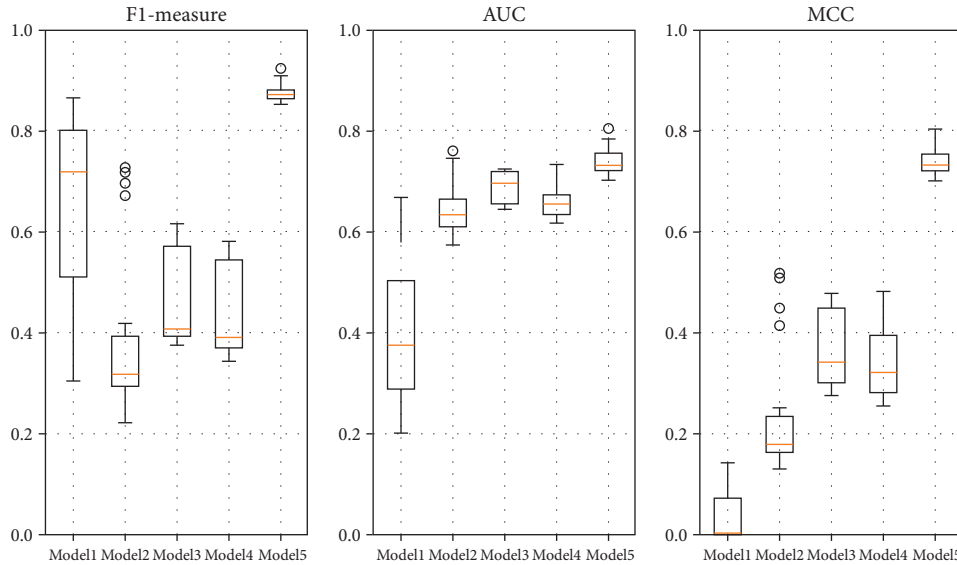


FIGURE 7: Boxplot of F1-measure, AUC, MCC of ALTRA (Model 1), TCA (Model 2), Peters filter (Model 3), Burak filter (Model 4), and TFIA (Model 5).

The attempt on more scenarios will be included in our future research plan.

## 6. Conclusion and Future Work

To tackle the problems of class imbalance and different data distribution among projects in cross-project defect prediction (CcrossPDP) which is the typical application of pattern recognition, we propose a model based on two-phase feature importance amplification (TFIA) in this paper. In the domain adaptation phase, TFIA reduces the differences in data distribution between the source and target projects by adding a certain percentage of samples from the target project to the source project. Meanwhile, correlation-based feature selection (CFS) with a greedy best-first search strategy for feature selection is utilized to obtain a feature subset with a weak feature-feature correlation and a strong feature-class correlation. The source and target project datasets filtered by this feature subset are input to the following classification phase. As the tree-based classifiers are sensitive to features with high feature-class correlation, Random Forest, the integrated method based on Decision Tree for pattern recognition, functions in the classification phase. Multiple decision trees can amplify the importance of features and help each other to improve the performance of the whole prediction model.

We conducted four experiments to validate TFIA on 20 project pairs from the widely adopted AEEEM dataset. The first experiment demonstrates that the domain adaptation approach we designed is effective and has a significant improvement on the overall model performance. The second experiment demonstrates that Random Forest outperforms other classifiers on the dataset processed by our domain adaptation approach. The third experiment analyses the

influence of the proportion of samples added from the target projects on the model. The experiment proves that adding 40% of the target item data makes great improvement on the model performance considering the practical application scenarios. In the fourth experiment, we compare TFIA with four other classical research methods, and the experimental result proves that TFIA provides better and more stable performance.

In this paper, our method has been proved to have an improvement effect on CPDP, but there are still some parts that need to be further studied. We will follow two aspects in our future research. The first is to validate the effectiveness of TFIA on other datasets and make it more scalable and robust since our experiments use the subprojects from the same root project. Secondly, since the parameters used in this paper are defaulted, which have a large impact on the prediction model, we will seek to find the effect of adjusting parameters on the model performance. We will also look for the most suitable parameters and the methods that can automatically adjust the parameters according to the real application scenarios.

## Data Availability

A publicly available dataset is used for this study (AEEEM).

## Conflicts of Interest

The authors declare no conflicts of interest.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (no. 61702044).

## References

- [1] R. S. Wahono, "A systematic literature review of software defect prediction," *Journal of Software Engineering*, vol. 1, no. 1, pp. 1–16, 2015.
- [2] N. Kalaivani and R. Beena, "Overview of software defect prediction using machine learning algorithms," *International Journal of Pure and Applied Mathematics*, vol. 118, no. 20, pp. 3863–3873, 2018.
- [3] Y. Kamei and E. Shihab, "Defect prediction: accomplishments and future challenges," in *Proceedings of the IEEE 23rd international conference on software analysis, evolution, and reengineering (SANER)*, vol. 5, March 2016.
- [4] D. Gray, D. Bowes, N. Davey, and Y. Sun, "The misuse of the NASA metrics data program data sets for automated software defect prediction," in *Proceedings of the 15th Annual Conference on Evaluation & Assessment in Software Engineering (EASE 2011)*, April 2011.
- [5] N. E. Fenton and M. Neil, "A critique of software defect prediction models," *IEEE Transactions on Software Engineering*, vol. 25, no. 5, pp. 675–689, 1999.
- [6] M. Cheng, G. Wu, M. Jiang, H. Wan, G. You, and Yuan, "Heterogeneous defect prediction via exploiting correlation subspace," *SEKE*, pp. 171–176, 2016.
- [7] T. Menzies, J. Greenwald, and A. Frank, "Data mining static code attributes to learn defect predictors," *IEEE Transactions on Software Engineering*, vol. 33, no. 1, pp. 2–13, 2006.
- [8] Z. He, F. Shu, Y. Yang, M. Li, and Q. Wang, "An investigation on the feasibility of cross-project defect prediction," *Automated Software Engineering*, vol. 19, no. 2, pp. 167–199, 2012.
- [9] S. Hosseini, B. Turhan, and D. Gunarathna, "A systematic literature review and meta-analysis on cross project defect prediction," *IEEE Transactions on Software Engineering*, vol. 45, no. 2, pp. 111–147, 2017.
- [10] J. Nam, S. J. Pan, and S. Kim, "Transfer Defect learning," in *Proceedings of the 2013 35th International Conference on Software Engineering (ICSE)*, 26 September 2013.
- [11] D. Ryu, J. Jang, and J. Baik, "A hybrid instance selection using nearest-neighbor for cross-project defect prediction," *Journal of Computer Science and Technology*, vol. 30, no. 5, pp. 969–980, 2015.
- [12] D. Ryu, J.-I. Jang, and J. Baik, "A transfer cost-sensitive boosting approach for cross-project defect prediction," *Software Quality Journal*, vol. 25, no. 1, pp. 235–272, 2017.
- [13] S. Hosseini, B. Turhan, and M. Mäntylä, "A benchmark study on the effectiveness of search-based data selection and feature selection for cross project defect prediction," *Information and Software Technology*, vol. 95, pp. 296–312, 2018.
- [14] A. O. Balogun, S. Basri, J. A. Said, V. E. Adegemo, and A. A. Imam, "Software Defect Prediction," *Analysis of Class Imbalance and Performance Stability*, vol. 14, no. 1, 2019.
- [15] Y. Jiang, M. Li, and Z. Zhou, "Software defect detection with rocus," *Journal of Computer Science and Technology*, vol. 26, no. 2, pp. 328–342, 2011.
- [16] D. Ryu, O. Choi, and J. Baik, "Value-cognitive boosting with a support vector machine for cross-project defect prediction," *Empirical Software Engineering*, vol. 21, pp. 43–71, 2016.
- [17] N. Japkowicz and S. Stephen, "The class imbalance problem: a systematic study1," *Intelligent Data Analysis*, vol. 6, no. 5, pp. 429–449, 2002.
- [18] V. Svetnik, A. Liaw, C. Tong, J. C. Culberson, R. P. Sheridan, and B. P. Feuston, "Random forest: a classification and regression tool for compound classification and QSAR modeling," *Journal of Chemical Information and Computer Sciences*, vol. 43, no. 6, pp. 1947–1958, 2003.
- [19] G. Biau and E. Scornet, "A random forest guided tour," *Test*, vol. 25, no. 2, pp. 197–227, 2016.
- [20] D. Bowes, T. Hall, and J. Petrić, "Software defect prediction: do different classifiers find the same defects?" *Software Quality Journal*, vol. 26, no. 2, pp. 525–552, 2018.
- [21] B. Turhan, "On the relative value of cross-company and within-company data for defect prediction," *Empirical Software Engineering*, vol. 14, no. 5, pp. 540–578, 2009.
- [22] F. Peters, T. Menzies, and A. Marcus, "Better cross company defect prediction," in *Proceedings of the 10th Working Conference on Mining Software Repositories (MSR)*, May 2013.
- [23] S. J. Pan and I. W. Tsang, "Domain adaptation via transfer component analysis," *IEEE Transactions on Neural Networks*, vol. 22, no. 2, pp. 199–210, 2010.
- [24] P. He, "An improved method for cross-project defect prediction by simplifying training data," *Mathematical Problems in Engineering*, vol. 2018, Article ID 2650415, 2018.
- [25] Z. Sun, J. Lic, and H. Sund, "CFPS: collaborative filtering based source projects selection for cross-project defect prediction," *Applied Soft Computing*, vol. 99, Article ID 106940, 2021.
- [26] H. Alsawalqah, "Hybrid SMOTE-ensemble approach for software defect prediction," *Computer Science On-Line Conference*, Springer, Cham, 2017.
- [27] Z. Yuan, "ALTRA: cross-project software defect prediction via active learning and tradaboost," *IEEE Access*, vol. 8, pp. 30037–30049, 2020.
- [28] C. Liu, "A two-phase transfer learning model for cross-project defect prediction," *Information and Software Technology*, vol. 107, pp. 125–136, 2019.
- [29] C. Jin, "Cross-project software defect prediction based on domain adaptation learning and optimization," *Expert Systems with Applications*, vol. 171, Article ID 114637, 2021.
- [30] Y. Zhang and J. Sun, "An empirical study of classifier combination for cross-project defect prediction," in *Proceedings of the IEEE 39th Annual Computer Software and Applications Conference*, Taichung, Taiwan, July 2015.
- [31] L. Chen, B. Fang, and Z. Shang, "Negative samples reduction in cross-company software defects prediction," *Information and Software Technology*, vol. 62, pp. 67–77, 2015.
- [32] X. Yu, M. Wu, and Y. Jian, "Cross-company defect prediction via semi-supervised clustering-based data filtering and MSTR-based transfer learning," *Soft Computing*, vol. 22, no. 10, pp. 3461–3472, 2018.
- [33] K. Kira and Rendell, "A practical approach to feature selection," *Machine Learning Proceedings 1992*, Morgan Kaufmann, Japan, pp. 249–256, 1992.
- [34] M. A. Hall, "University of Waikato, Department of Computer Science," *Correlation-based Feature Selection for Machine Learning*, Hamilton, NewZealand, 1999.
- [35] F. Xie, M. Müller, and R. Holte, "Jasper: the art of exploration in greedy best first search," *The Eighth International Planning Competition*, pp. 39–42, IPC-2014, Edmonton, Canada, 2014.
- [36] A. Karegowda, A. Manjunath, and M. Jayaram, "Comparative study of attribute selection using gain ratio and correlation based feature selection," *International Journal of Information Technology and Knowledge Management*, vol. 2, no. 2, pp. 271–277, 2010.
- [37] M. A. Hall, "University of Waikato, Department of Computer Science." *Correlation-based Feature Selection of Discrete and Numeric Class Machine Learning*, New Zealand, 2000.

- [38] Z. Zhou, J. Yang, Y. Qi, and Y. Cai, "Support vector machine and back propagation neural network approaches for trip mode prediction using mobile phone data," *IET Intelligent Transport Systems*, vol. 12, no. 10, pp. 1220–1226, 2018.
- [39] F. Zhang, Q. Zheng, Y. Zou, and A. Hassan, "Cross-project defect prediction using a connectivity-based unsupervised classifier," in *Proceedings of the IEEE/ACM 38th International Conference on Software Engineering (ICSE)*, IEEE, May 2016.
- [40] T. Pushphavathi, V. Suma, and V. Ramaswamy, "A novel method for software defect prediction: hybrid of FCM and random forest," in *Proceedings of the International Conference on Electronics and Communication Systems (ICECS)*, Feb2014.
- [41] M. D'Ambros, M. Lanza, and R. Robbes, "An extensive comparison of bug prediction approaches," in *7th IEEE Working Conference on Mining Software Repositories (MSR 2010)*, May 2010.
- [42] I. H. Witten and E. Frank, "Data mining," *Acm Sigmod Record*, vol. 31, no. 1, pp. 76-77, 2002.
- [43] Wikipedia contributors, "Matthews correlation coefficient," *Wikipedia, the Free Encyclopedia*. *Wikipedia, the Free Encyclopedia*, 2021.
- [44] S. Boughorbel, F. Jarray, and M. ElAnbari, "Optimal classifier for imbalanced data using Matthews Correlation Coefficient metric," *PLoS One*, vol. 12, Article ID e0177678, 2017.



## Research Article

# Combined Economic Emission Dispatch of Microgrid with the Incorporation of Renewable Energy Sources Using Improved Mayfly Optimization Algorithm

Karthik Nagarajan <sup>1</sup>, Arul Rajagopalan <sup>2</sup>, S. Angalaeswari <sup>2</sup>, L. Natrayan <sup>3</sup>,  
and Wubishet Degife Mammo <sup>4</sup>

<sup>1</sup>Department of Electrical & Electronics Engineering, Hindustan Institute of Technology & Science, Chennai, Tamil Nadu, India

<sup>2</sup>School of Electrical Engineering, Vellore Institute of Technology, Chennai 600127, Tamil Nadu, India

<sup>3</sup>Department of Mechanical Engineering, Saveetha School of Engineering, SIMATS, Chennai 602105, Tamil Nadu, India

<sup>4</sup>Mechanical Engineering Department, Wollo University, Kombolcha Institute of Technology, Kombolcha, South Wollo-208, Amhara, Ethiopia

Correspondence should be addressed to Arul Rajagopalan; arulphd@yahoo.co.in and Wubishet Degife Mammo; wubishetdegife7@gmail.com

Received 24 October 2021; Revised 26 November 2021; Accepted 24 March 2022; Published 18 April 2022

Academic Editor: Daqing Gong

Copyright © 2022 Karthik Nagarajan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Electricity can be provided to small-scale communities like commercial areas and villages through microgrid, one of the small-scale, advanced, and independent electricity systems out of the grid. Microgrid is an appropriate choice for specific purposes reducing emission and generation cost and increasing efficiency, reliability, and the utilization of renewable energy sources. The main objective of this paper is to elucidate the combined economic emission dispatch CEED problem in the microgrid to attain optimal generation cost. A combined cost optimization approach is examined to minimize operational cost and emission levels while satisfying the load demand of the microgrid. With this background, the authors proposed a novel improved mayfly algorithm incorporating Levy flight to resolve the combined economic emission dispatch problem encountered in microgrids. The islanded mode microgrid test system considered in this study comprises thermal power, solar-powered, and wind power generating units. The simulation results were considered for 24 hours with varying power demands. The minimization of total cost and emission is attained for four different scenarios. Optimization results obtained for all scenarios using IMA give a comparatively better reduction in system cost than MA and other optimization algorithms considered revealing the efficacy of IMA taken for comparison with the same data. The proposed IMA algorithm can solve the CEED problem in a grid-connected microgrid.

## 1. Introduction

Microgrid is one of the advanced small-scale centralized electricity systems and it usually contains energy storage resources, Distributed Generation (DG) units, and loads. Microgrids are generally designed and installed nearby energy consumers in a confined community [1]. But there has been a drastic increase in recent years regarding installing renewable energy sources in microgrids, owing to environmental advantages and low cost compared to its counterparts [2]. Microgrid meets different load

requirements in residential, commercial, agriculture, and industrial sectors [3]. Microgrids can function under two different modes: grid-connected and islanded modes. In the former mode, the microgrid and main grid are connected, while in the latter the microgrid is isolated from the main grid during an emergency outbreak. It continues its power delivery functions to local loads as usual [4]. There are loads of advantages present in using microgrid reduction of carbon emission and generation cost, thanks to renewable energy sources, and increased reliability and power quality, etc. [5].



Optimal operations and effective planning of electric power generation systems are the two most crucial elements in electric industries. Controlling and operating power systems, cost-efficient load dispatch (Economic Load Dispatch, i.e., ELD) related problems are too much concerning to address [6]. Power system optimization problems that employ ELD are useful in identifying the most suitable, cheap, and seamless operations with the regulation of outputs produced by different power generation units that meet the load demands. ELD has a primary aim to mitigate the overall cost incurred upon power generation without compromising or producing any constraints [7]. ELD determines active power output generated by different power generation systems to attain the objective functions and simultaneously overcome many problems [8]. It is important to develop novel power management algorithms to translate the microgrid as a viable and happening choice compared to traditional power systems [9]. The mitigation of both power generation costs and the emission of environmental pollutants is the sole aim of utility operators. For these goals, two contradictory objectives must be considered. Combined economic emission dispatch (CEED) is utilized to mitigate emission levels from all generating units and costs incurred by the operating units [9].

CEED problem has been discussed earlier by different authors who proposed several optimization techniques to overcome the issue [10]. In [10], the researchers proposed a balanced trade-off method to resolve the ECED (Environment Constrained Economic Dispatch) problem. The study conducted a first-of-its-kind comparative analysis of three methods: Fractional Programming (FP), ECED, and price penalty factor (PPF) to overcome the CEED problem. Chimp Optimization Algorithm (ChOA) was proposed to address the optimal design of microgrid which comprises PV panels, wind turbines, and battery storage systems [11]. Optimization results had been compared with Improved Grey Wolf Optimizer (IGWO) and Grey Wolf Optimizer (GWO). Optimal sizing of photovoltaic cell and solar water heater by considering environmental parameters and fuel-saving was carried out in [12]. Energy and economic analysis of solar energy-based cogeneration system for a building in Saveh City were studied. The researchers simulated the model in a 3-Unit dynamic test system incorporating renewable energy sources. In the study conducted by Alammoush [13], Bernstein-search differential evolution (BSDE) algorithm was proposed to resolve the Dynamic Combined Heat and Power Economic emission Dispatch (DCHPED) generation problem in microgrid comprised of renewable energy sources, fixed nondeferrable and deferrable loads, fossil-fuel combined heat and power units, and thermal energy storage devices. In [14], whale optimization algorithm (WOA) has been applied to carry out the combined economic emission dispatch problem. The simulation was performed by considering four different load sharing scenarios among the distributed energy resources. The researchers also conducted ANOVA and Wilcoxon signed-rank tests to validate the supreme characteristic of WOA. In [15], Stochastic Fractal Search (SFS) algorithm was applied to resolve multiobjective economic emission dispatch

problems that arise in combined heat and power (CHP) generation. This study was conducted in large microgrids in which solar-powered generating units, wind power units, and fossil-fuel-powered generating units were installed.

Collective Neuro Dynamic Optimization (CNO) method was proposed in the earlier study. In this study, the authors combined a heuristic approach and projection neural network (PNN) to optimize the scheduling of an electrical microgrid containing ten thermal generators and mitigate the costs incurred upon emission and generation [16]. A mixed-integer nonlinear programming formulation was proposed in [17] to dispatch the distributed generators cohesively. Further, the model was also aimed at fulfilling the water demands and the building's thermal energy requirements in a standalone water energy microgrid.

In [18], Giza Pyramids Construction (GPC) was proposed to implement the optimal design of an isolated microgrid. Net present cost (NPC), Levelized Cost of Energy (LCOE), loss of power supply probability, and availability index were considered objective functions. Modified adaptive accelerated particle swarm optimization (MAACPSO) algorithm was proposed to investigate the grid-tied PV systems reliability [19]. This study focused on the probability analysis and reliability assessment of the components of grid-tied PV systems through IEEE 24 bus integrated with PV system with four different case studies. A first-of-its-kind Sequential Optimization Strategy (SOS) was formulated in the study conducted earlier to allocate active and reactive power to Dispatchable Distributed Generator (DDG) units in an optimal manner. These DDG units are installed in a droop-controlled islanded AC microgrid [20]. The research proposed improved Quantum Particle Swarm Optimization (QPSO) earlier [21] to address the Short-Term Economic Environmental Dispatch (EED) problem in a microgrid. Recently, a multiobjective seeker optimization algorithm has been proposed to analyze the influence of charging and discharging behaviour of electric vehicles and demand side response resources on the economic functioning of PV-connected microgrid systems. The model considered three objectives: power fluctuation between microgrid and main grid, comprehensive operating cost of the microgrid, and utilization rate of photovoltaic energy [22].

In [23], the researchers proposed a mathematical optimization approach to achieve an optimal operation upon economic dispatch in DC microgrid. To allot the schedules for unit commitment and achieve economic dispatch in microgrid, the study conducted earlier [24] proposed an enhanced real-coded genetic algorithm in the enhanced mixed-integer linear programming (MILP) based method. The authors proposed a stochastic model in [25] to manage CHP-based microgrids optimally. The study took economic, reliability, and environmental aspects into consideration. Nonconvex and nonlinear stochastic problems are used to resolve complexity. The Exchange Market Algorithm (EMA) was proposed in a study. This study considered three contradictory objectives through a weighted sum approach to resolving the multiobjective problem as a single-objective problem.

A multiobjective optimal dispatch model was proposed in [26] for a grid-connected microgrid. In this study, the authors considered reducing environmental protection costs and the generation cost of the microgrid with the incorporation of an enhanced PSO algorithm. In [27], a genetic algorithm is applied to optimize the focal area of the parabolic trough concentration photovoltaic/thermal system. The objective function was considered a combination of electrical efficiency and thermal efficiency. A parabolic trough concentrating photovoltaic thermal (CPVT) was utilized to afford the energy required for a residential building [28]. CPVT was used as a source of heat and cooling and electrical energy for the building and simulation was carried out using TRNSYS software. Samy et al. addressed the problem of power outages in distant districts by taking advantage of the available renewable energy resources in the contiguous environment [29]. Hybrid Firefly and Harmony Search optimization technique (HFA/HS) was implemented to improve the net present cost of the proposed hybrid system which comprises photovoltaic (PV), wind turbine (WT), and fuel cell (FC). MINLP, a novel optimization model, was proposed in [30] to resolve the economic dispatch problem of a microgrid which contains numerous units of wind turbines (WTs), heat-only units, traditional power generators, photovoltaic (PV) systems, CHP units, and battery storage systems under certain uncertainties. In the study conducted earlier [31], MBGSA (Memory-Based Gravitational Search Algorithm) was proposed to overcome the ELD problem. In one of the research investigations [32], a novel Multiobjective Virus Colony Search (MOVCS) was proposed to elucidate the multiobjective dynamic economic emission dispatch (DEED) problem. This model aims to mitigate the emissions produced by fossil-fuel power generators and simultaneously reduce the cost incurred upon wind-thermal electrical energy costs. In [33], HOMER software is used to model and simulate a wind-solar hybrid system independent of the national grid in the northwest of Iran. A multiobjective particle swarm optimization technique is proposed to optimize the sizing of a green energy system connected to a randomly disrupted grid [34]. The energy cost for evaluating hybrid system economies, the loss of probability of power supply (LPSP) for reliability assessments, and the System Surplus Energy Rates (SSER) were considered objective functions for evaluating hybrid system compatibility and efficiency. A sustainable energy distribution configuration for microgrids integrated into the national grid using back-to-back converters in a renewable power system was examined in [35]. Different scenarios of several sustainability schemes of power management in microgrids were analyzed.

SGEO (Social Group Entropy Optimization) technique was proposed in [36] to resolve Fuel Constrained Dynamic Economic Dispatch (FCDED) with Demand Side Management (DSM). The technique combined the pumped hydrostorage plant with renewable energy sources. This research used a stochastic fractal search algorithm to overcome the biobjective combined heat and power economic dispatch (CHPED) problem [37].

In [38], mayfly algorithm (MA) was proposed by Dr. Konstantinos Zervoudakis in 2021. In this paper, an improved mayfly optimization algorithm was investigated with the help of a microgrid model under varying scenarios. The results were contrasted against recent state-of-the-art algorithms that also employed the same microgrid model.

The contributions of the current research paper are summarized herewith.

- (i) An improved version of the mayfly optimization algorithm incorporating Levy flight is proposed to elucidate the microgrid test system's CEED problem.
- (ii) An improved mayfly optimization algorithm is proposed in addition to Levy flight to overcome the CEED problem encountered in microgrid test system.
- (iii) Levy flight has been leveraged in this study since it possesses huge advantages in not engaging local optimal. An optimal trade-off is provided by the proposed algorithm between exploration and exploitation phases.
- (iv) The authors validated the supremacy of the proposed algorithm in terms of resolving the CEED problem under two different objective functions that involve advanced energy sources.
- (v) Compared with existing population-based optimization tools such as PSO and GA, only a few control parameters exist in IMA. This feature helps in making it the best optimization procedure. IMA-based CEED was authenticated as a unique and robust technique since it incurred less total generation cost than the solution even after conducting multiple random trials.

## 2. Mathematical Formulation of CEED for Microgrid

*2.1. Combined Economic Emission Dispatch (CEED).* The simultaneous mitigation of economic and environmental dispatch objective functions remains the primary objective for the CEED problem in the microgrid. In other terms, the total fuel cost must be reduced, while at the same time the emission levels should also be mitigated without compromising the constraints. So, it is suggested to formulate CEED as a single optimization problem as given herewith [9].

$$\min \text{TC} = \sum_{i=1}^{\text{NG}} \{F_i(P_{Gi}), E_i(P_{Gi})\}. \quad (1)$$

Here, TC denotes the total operating cost that should be minimized. The fuel cost of the  $i^{\text{th}}$  generator is represented as  $F_i(P_{Gi})$ , the emission level of the  $i^{\text{th}}$  generator is represented as  $E_i(P_{Gi})$ ,  $P_{Gi}$  denotes the  $i^{\text{th}}$  generating unit's output power, and finally, NG corresponds to the whole generating unit count.

**2.1.1. Minimization of Fuel Cost.** In general, the fuel cost function is denoted through the quadratic equation given below [9]:

$$F_t = \sum_{i=1}^{NG} F_i(P_{Gi}) = \sum_{i=1}^{NG} (x_i + y_i P_{Gi} + z_i P_{Gi}^2). \quad (2)$$

Here,  $F_t$  corresponds to overall fuel cost incurred in terms of \$ and  $x_i$  (\$/h),  $y_i$  (\$/MWh), and  $z_i$  (\$/MW<sup>2</sup>h) correspond to the cost coefficients of  $i^{\text{th}}$  generating unit.

**2.1.2. Minimization of Emission.** When fossil fuels are used, the generators emit different sorts of pollutants. So, pollution mitigation forms the primary goal of most power system operations. Equation (3) denotes the expression for total emission from [5, 9]:

$$E_t = \sum_{i=1}^{NG} E_i(P_{Gi}) = \sum_{i=1}^{NG} (\alpha_i + \beta_i P_{Gi} + \gamma_i P_{Gi}^2). \quad (3)$$

Here,  $E_t$  denotes the overall emission and  $\alpha_i$  (kg/h),  $\beta_i$  (kg/(MWh)),  $\gamma_i$  (kg/(MW<sup>2</sup>h)) correspond to the  $i^{\text{th}}$  generating unit's emission coefficients.

**2.1.3. Total Generation Cost of CEED Problem.** It is possible to convert the dual-objective optimization problem, focusing emission, and fuel cost, into a single-objective optimization problem with the induction of PPF (price penalty factor) as given earlier [9]:

$$\text{minTC} = F_t + \Lambda \times E_t. \quad (4)$$

Here,  $\Lambda$  denotes the price penalty factor (PPF), which is calculated as a ratio between fuel cost and the emission of the corresponding generating unit (\$/kg). PPFs are of different types, while in the current study, the authors use min-max types sourced from [5, 9] for comparison. Following is the equation for min-max type [9]:

$$\Lambda_i = \frac{F_t(P_{Gi}^{\min})}{E_t(P_{Gi}^{\max})}, \quad i = 1, 2, \dots, NG, \quad (5)$$

where  $P_{Gi}^{\max}$  and  $P_{Gi}^{\min}$  Here,  $\Lambda$  denotes the ratio between maximum fuel cost and maximum emission of the corresponding generator in \$/kg [9]. The maximum and minimum output power of the generator combines emission with fuel cost. Afterwards, TC corresponds to the total operating cost of \$.

The following is the list of steps to be followed to determine the price penalty factor for a specific load demand [39].

- (i) The ratio between minimum fuel cost and the maximum emission of every generating unit should be determined.
- (ii) Price Penalty Factor values are sorted out in ascending order.

(iii) The maximum capacity of every unit ( $P_{Gi}^{\max}$ ) one is added at a time that starts from the lowest  $\Lambda_i$ , until  $\sum P_{Gi}^{\max} \geq P_D$ .

(iv) Then,  $\Lambda_i$ , which has an association with the lowest unit in this process, remains the tentative PPF value ( $\Lambda$ ) for the load under consideration.

So, a modified PPF ( $\Lambda$ ) is utilized to arrive at the exact value for specific load demand based on the interpolation of  $\Lambda$  values corresponding to their load demand values.

**2.2. Cost Functions of Renewable Energy Sources.** Across the globe, renewable energy sources are the foremost choice of transmission when energy is produced, compared to traditional generators. In such a scenario, solar and wind power can be denoted as negative loads and can be used to mitigate the total load demand in the system [9]. However, the economic dispatch solution is considered a base to distribute the rest of the load demands on traditional generators. In the current study, the CEED solution for the microgrid takes cost functions of wind and solar-powered generating units into account. From [9], the input data for cost and emission coefficients are considered.

**2.2.1. Cost Function of Wind Power Generating Unit.** Current economic analysis is conducted for wind-based power generation and the specific cost can be determined with inputs, operation, maintenance, and equipment costs. This cost function is expressed as per [5, 9]:

$$C_w(P_w) = \left( \frac{r}{[1 - (1 + r)^{-N}]^P} + O^E \right) P_w. \quad (6)$$

Here,  $P_w$  denotes the wind power produced in terms of kW,  $r$  corresponds to the interest rate,  $a$  denotes the Annuitization coefficient,  $N$  corresponds to lifetime investment in terms of years, and  $l^P$  and  $O^E$  correspond to the costs incurred upon investment per unit installed power (\$/kW) and operating and maintenance costs per unit installed power (\$/kW), respectively.

The 24-hour data for the wind power generating unit is considered from [5]. The parameters required for wind power cost function are chosen from [5, 9].

**2.2.2. Cost Function of Solar Power Generating Unit.** Similar to wind power, solar-powered generating unit's cost function is expressed as in [5, 9]:

$$C_s(P_s) = \left( \frac{r}{[1 - (1 + r)^{-N}]^P} + O^E \right) P_s, \quad (7)$$

where  $P_s$  is the output power from the solar-powered generating unit,  $r$  corresponds to interest rate,  $N$  denotes lifetime investment in terms of years, and  $l^P$  and  $O^E$  correspond to investment costs made upon per unit installed power (\$/kW) and operating and maintenance costs per unit installed power (\$/kW), respectively.

The 24-hour data of the solar power generating unit is considered [5]. The parameters required for solar power cost function are chosen from [5, 9].

$$C_T = F_t + \Lambda \times E_t + \left( \frac{r}{[1 - (1 + r)^{-N}]^P} + O^E \right) P_w + \left( \frac{r}{[1 - (1 + r)^{-N}]^P} + O^E \right) P_s. \quad (8)$$

**2.4. Constraints.** The researcher considered both generator capacity and power balance constraints to contrast with existing optimization algorithms.

**2.4.1. Islanded Mode of Microgrid.** The current study considered islanded mode microgrid to compare with the optimization results achieved in [5, 9]. There is no trade-off for the power between the main grid and the microgrid in this mode. So, the microgrid needs to fulfil the local or confined community load demands.

**2.4.2. Power Balance Constraint.** The load demand must be equal to that of the total power generation [9].

$$\sum_{i=1}^{NG} P_{Gi} + P_w + P_s = P_D. \quad (9)$$

Here,  $P_D$  corresponds to the total load demand.

**2.4.3. Generation Capacity Constraints.** Every generating unit's output power gets flanked by both lower and upper bounds [9].

$$P_{Gi}^{\min} \leq P_{Gi} \leq P_{Gi}^{\max}. \quad (10)$$

Here,  $P_{Gi}^{\min}$  and  $P_{Gi}^{\max}$  correspond to the minimum and maximum output powers of the  $i^{\text{th}}$  generating unit correspondingly.

### 3. Improved Mayfly Algorithm

Mayfly algorithm takes its inspiration from the social behaviour of mayflies, especially how they mate with each other [38]. It is assumed that mayflies are instantly considered adults as soon as the eggs are hatched. Leaving beside the period of their life, only the fittest mayflies tend to survive. Each mayfly has a position in search space that corresponds to a solution that overcomes the problem. RAND functions are utilized in conventional mayfly algorithm to produce novel variables that lead to local optimal. To increase MA's searching ability and create an optimal solution, the researchers integrated MA with Levy flight. If a Levy flight-based approach is utilized for system identification, it achieves rapid convergence and does not entail derivative information [40], attributed to stochastic random search, in line with the Levy flight concept [41]. Levy flight contributes heavily to increasing the optimal solution's local

**2.3. Total Cost of CEED for Microgrid.** The equation for the total CEED cost for the microgrid is shown below. This value gets minimized based on the cost functions regarding wind and solar power [9].

search avoidance and local trapping [42]. The flowchart of the proposed IMA algorithm is shown in Figure 1.

The steps required for the proposed mayfly optimization algorithm works are described as follows:

**Step 1.** Two mayfly sets, each representing a male and a female population, should be generated randomly. Then, every mayfly is arbitrarily placed in problem space as a candidate solution which is denoted by a  $d$ -dimensional vector  $P_{Gi} = (P_{G1}, \dots, P_{Gd})$ . Then the performance is assessed based on the predefined objective function  $f(C_T(P_{Gi}))$ .

**Step 2.** A mayfly's velocity  $\mathbf{v} = (v_1, \dots, v_d)$  is initialized through its positional change. Its direction is decided as a hybrid interaction between individuals and the social flying experiences. To be specific, every mayfly tends to alter its trajectory in alignment with its personal best position (pbest) so far. It also alters based on the best position achieved by any other mayfly present in the swarm so far (gbest).

**Step 3.** The population of the male mayflies is initialized as  $P_{Gmi}$  ( $i = 1, 2, \dots, NG$ ) with velocities  $v_{mi}$ . The male mayflies, gathered in swarms, denote that the position of every mayfly gets altered in alignment with its individual's experience and that of the neighbor's.  $P_{Gi}^t$  is assumed to be the current position of mayfly  $i$  in search space at time step  $t$  and the position gets altered with the addition of velocity  $v_i^{t+1}$ , to the current position. This notation is formulated as given herewith.

$$P_{Gmi}^{t+1} = P_{Gmi}^t + v_i^{t+1}. \quad (11)$$

Male mayflies are considered as present a few meters above the water, with  $P_{Gim}^0 U(P_{Gmmin}, P_{Gmmax})$ , performing nuptial dance. It can be assumed that these mayflies lack great speeds due to constant movements. This results in the calculation of a male mayfly's velocity  $i$  as follows [38]:

$$v_{ij}^{t+1} = g * v_{ij}^t + a_1 e^{-\beta r_p^2} (pbest_{ij} - P_{Gmij}^t) + a_2 e^{-\beta r_g^2} (gbest_j - P_{Gmij}^t). \quad (12)$$

Here,  $v_{ij}^t$  corresponds to mayfly  $i$ 's velocity in dimension  $j = 1, \dots, n$  at time step  $t$ ,  $P_{Gmij}^t$  denotes the mayfly's  $i^{\text{th}}$  position in dimension  $j$  at time step  $t$ ,  $a_1$  and  $a_1$  correspond to positive attraction constants utilized in scaling up the contribution of cognitive and social components,



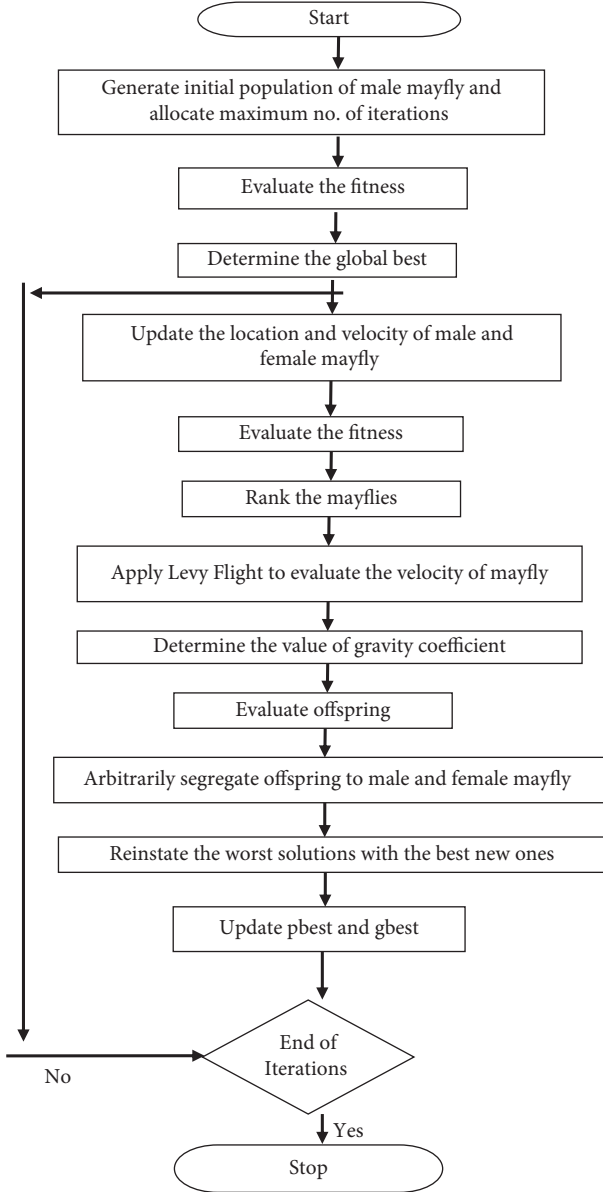


FIGURE 1: Flowchart of the proposed improved mayfly optimization algorithm.

respectively. Furthermore,  $\text{pbest}_i$  denotes the mayfly  $i^{\text{th}}$  best position which it had ever visited. Based on the minimization problems under consideration, the personal best position  $\text{pbest}_{ij}$  at the next time step  $t + 1$  is determined as given herewith.

$$\text{pbest}_i = \begin{cases} P_{\text{Gmi}}^{t+1}, & \text{iff } (P_{\text{Gmi}}^{t+1}) < f(\text{pbest}_i) \\ \text{is kept the same,} & \text{otherwise} \end{cases} \quad (13)$$

Following is the equation for the global best position  $\text{gbest}$  at time step  $t$ .

$$\begin{aligned} \text{gbest} &= \{ \text{pbest}_1, \text{pbest}_2, \dots, \text{pbest}_N, |f(\text{cbest}) \}, \\ &= \min \{ f(\text{pbest}_1), f(\text{pbest}_2), \dots, f(\text{pbest}_{\text{NG}}) \}. \end{aligned} \quad (14)$$

Here  $\beta$  represents the fixed visibility coefficient used in (7). It is utilized to confine the visibility of the mayfly to others. Further,  $r_p$  denotes the Cartesian distance between  $P_{\text{Gi}}$  and  $\text{pbest}_i$  and  $r_g$  corresponds to the Cartesian distance between  $P_{\text{Gi}}$  and  $\text{gbest}$ . Following is the equation used to determine these distances.

$$\|P_{\text{Gmi}} - X_i\| = \sqrt{\sum_{j=1}^n (P_{\text{Gmij}} - X_{ij})^2}, \quad (15)$$

where  $P_{\text{Gmij}}$  corresponds to the  $j^{\text{th}}$  element of mayfly  $i$  and  $X_i$  denotes the  $\text{pbest}_i$  or  $\text{gbest}$ . If the algorithm needs to function appropriately, then the best mayflies present in the swarm must continuously perform the up-and-down nuptial dance. So, the velocity of these best mayflies must be kept on changing which is calculated as follows [38]:

$$v_{ij}^{t+1} = v_{ij}^t + d \times r. \quad (16)$$

Here,  $d$  denotes the coefficient of nuptial dance whereas the random value in the range of  $[-1, 1]$  is denoted by  $r$ .

*Step 4.* In this step, the female mayfly population is initialized  $P_{\text{Gfi}} (i = 1, 2, \dots, \text{NG})$  with velocities  $v_{fi}$ . Female mayflies tend not to gather as a swarm alike males. Instead, it tends to fly towards its male counterparts for mating.  $P_{\text{Gfi}}^t$  is assumed as the current position of female mayfly  $i$  in search space at time step  $t$ , while its position gets altered with the addition of velocity  $v_i^{t+1}$  to the current position, i.e.,

$$P_{\text{Gfi}}^{t+1} = P_{\text{Gfi}}^t + v_i^{t+1}. \quad (17)$$

Here, due to  $P_{\text{Gfi}}^0 U(P_{\text{Gfmin}}, P_{\text{Gfmax}})$  one cannot randomize the attraction process. So, the model is decided to be a deterministic process. As a result, their velocities are determined as given herewith in the presence of minimization problems [38].

$$v_{ij}^{t+1} = \begin{cases} g * v_{ij}^t + a_2 e^{-\beta r_{\text{mf}}} (P_{\text{Gmij}}^t - P_{\text{Gfi}}^t), & \text{if } f(P_{\text{Gfi}}) > f(P_{\text{Gmi}}), \\ g * v_{ij}^t + fl \times r, & \text{if } f(P_{\text{Gfi}}) \leq f(P_{\text{Gmi}}). \end{cases} \quad (18)$$

Here,  $v_{ij}^t$  corresponds to the female mayfly's velocity  $i$  in dimension  $j = 1, \dots, n$  at time step  $t$ ,  $P_{\text{Gfij}}^t$  denotes the female mayfly  $i$ 's position in dimension  $j$  at time step  $t$ , and  $a_2$  denotes the positive attraction constant whereas it remains a fixed visibility coefficient. Further, the gravity coefficient is denoted by  $g$ , and  $r_{\text{mf}}$  corresponds to the Cartesian distance between male and female mayflies. Here  $fl$  corresponds to a random walk coefficient and  $r$  denotes the random value in the range of  $[-1, 1]$ . This value is determined based on (15).

*Step 5.* In this step, the Levy flight approach is involved in calculating the velocity of a mayfly candidate solution. Equation (19) is used to determine the velocity of the mayfly candidate solution [38].

$$v_{ij}^{t+1} = \begin{cases} V_{\max}, & \text{if } v_{ij}^{t+1} > V_{\max}, \\ -V_{\max}, & \text{if } v_{ij}^{t+1} < -V_{\max}. \end{cases} \quad (19)$$

This stage uses the Levy flight approach to alter the position of the global finest component. Though the Levy flight method has been used for exploration purposes so far, it is associated with a specific search.

Here,  $V_{\max}$  is calculated as follows:

$$V_{\max} = \text{Levy}(\lambda) * (P_{G\max} - P_{G\min}). \quad (20)$$

Here,  $\delta$  corresponds to a scale factor designed in alignment with the search space element. The author fixed  $\delta$  as 1.

$$\text{Levy}(\lambda) = 0.01 \frac{r_5 \sigma}{|r_6|^{1/\beta}} \quad (21)$$

Further,  $\sigma$  is calculated as follows [42]:

$$\sigma = \left[ \frac{\Gamma(1 + \lambda) \sin(\pi(\lambda/2))}{\Gamma((1 + \lambda)/2) \lambda [2^{(\lambda-1)/2}]} \right]^{1/\lambda}. \quad (22)$$

Here,  $\Gamma(x) = (x-1)!$ ,  $r_5$  corresponds to  $r_6$  indiscriminate numbers that lie in the range of [0, 1], and  $1 < \beta \leq 2$ , where there is a constant value, i.e., 1.5 incorporated for  $\beta$  in the current study [40–42].

$\text{Levy}(\lambda)$  denotes the step length, incorporated by Levy distribution with infinite variance and mean values with  $1 < \lambda < 3$ .  $\lambda$  corresponds to the distribution factor, whereas the gamma distribution function is denoted by  $\Gamma(\cdot)$ .

*Step 6.* Gravity coefficient value calculation [38]:

Gravity coefficient  $g$  value can be considered a fixed number that lies in (0, 1].

$$g = g_{\max} - \frac{g_{\max} - g_{\min}}{\text{iter}_{\max}} \times \text{iter}, \quad (23)$$

where  $g_{\max}$ ,  $g_{\min}$  correspond to maximum and minimum values which can be taken for the gravity coefficient, and  $\text{iter}$  denotes the algorithm's current iteration, whereas the maximum count of iterations is denoted by  $\text{iter}_{\max}$ .

*Step 7.* Mayflies are mated and the offspring are evaluated.

The mating process between the mayflies is discussed by the crossover operator as given herewith. From the male and female population, each one parent is selected through the same selection process, i.e., the attraction of females towards the males. Specifically, fitness function-based or random selection of the parents can be made. In terms of the fitness function, the best female mates with the best male, the second-best female with the second-best male, etc. This crossover results in two offsprings for which the formulation is given herewith [38]:

$$\begin{aligned} \text{offspring1} &= L \times \text{male} + (1 - L) \times \text{female}, \\ \text{offspring2} &= L \times \text{female} + (1 - L) \times \text{male}. \end{aligned} \quad (24)$$

Here, *male* denotes the male parent, and female corresponds to the female parent, while  $L$  is a random value within a specific range. The initial velocity of the offspring is fixed as zero.

*3.1. The Pseudocode of the Improved Mayfly Algorithm.* The pseudocode of the improved mayfly algorithm is devised as follows:

- (1) Formulate the objective function  $f(C_T(P_{Gi}))$ ,  $P_{Gi} = (P_{Gi1}, \dots, P_{Gid})^T$
- (2) Set the male mayfly population  $P_{Gmi}$  ( $i = 1, 2, \dots, NG$ ) and velocities  $v_{mi}$
- (3) Set the female mayfly population  $P_{Gfi}$  ( $i = 1, 2, \dots, NG$ ) and velocities  $v_{fi}$
- (4) Evaluate solutions
- (5) Determine global best gbest
- (6) Do While stopping criteria are not meet
- (7) Update velocities and solutions of males and females
- (8) Evaluate solutions
- (9) Rank the mayflies
- (10) Apply Levy flight approach to evaluate the velocity of a mayfly candidate solution
- (11) Determine the value of the gravity coefficient
- (12) Mate the mayflies
- (13) Evaluate offspring
- (14) Separate offspring to male and female randomly
- (15) Reinststate worst solutions with the best new ones
- (16) Update pbest and gbest
- (17) End While
- (18) Post-process results and visualization

## 4. Results and Discussion

The current study used a microgrid model with three conventional generators, solar, and wind units. One of the generators is a combined heat and power generator, whereas the other two conventional generators are synchronous. As per [9], three conventional generators and their daily load profile details were used in the current study. During the improved mayfly optimization algorithm implementation, various parameters were chosen for the optimal search process.

In order to assess the proposed IMA, various scenarios were considered as given herewith.

- (i) All sources included
- (ii) Thermal power generating units without renewable sources
- (iii) Thermal power generating units with wind source only
- (iv) Thermal power generating units with solar source only

TABLE 1: Optimal generation schedule of microgrid for case 1.

Time (h)	RGM cost (\$/h) [5]	ACO cost (\$/h) [5]	CSA cost (\$/h) [5]	ISA cost (\$/h) [5]	IHS cost (\$/h) [9]	IAHS cost (\$/h) [9]	MHS cost (\$/h) [9]	MA (\$/h)	IMA (\$/h)
1	8529	7250	7153	7153	7090.5	7058.0	6942.8	6857.4	6824.6
2	8648	7511	7203	7203	7151.1	7130.2	7010.3	6904.7	6891.4
3	8675	7704	7278	7278	7170.8	7151.5	7100.7	7015.6	6994.8
4	8795	7742	7280	7285	7159.6	7130.8	7049.6	7020.3	7004.8
5	8758	8211	7545	7545	7528.2	7450.1	7377.2	7334.1	7309.7
6	8848	8459	7723	7679	7600.1	7572.2	7553.3	7544.2	7537.6
7	8964	8406	7457	7457	7444.2	7423.8	7294.1	7207.3	7189.5
8	9308	7923	7138	7138	7051.0	7050.3	6935.6	6879.7	6851.3
9	9609	9040	7731	7731	7660.3	7640.9	7576.4	7528.1	7505.8
10	10049	9599	7920	7937	7851.5	7845.4	7770.8	7752.6	7731.2
11	11520	11184	9231	9231	9152.0	9150.0	9073.4	9025.3	9006.8
12	12098	11616	9470	9470	9394.3	9381.3	9314.3	9271.4	9253.7
13	10676	10320	8482	8482	8400.3	8374.4	8326.2	8297.6	8273.9
14	9982	9707	8186	8186	8135.4	8119.9	8025.4	7958.9	7929.4
15	9569	9351	8154	8159	8100.6	8090.5	7984.4	7903.2	7892.7
16	9030	8469	7622	7626	7550.5	7539.6	7457.9	7419.7	7401.4
17	8872	8189	7526	7525	7470.6	7440.2	7362.6	7305.8	7291.3
18	9273	9061	8132	8131	8050.8	8040.4	7956.6	7904.3	7889.5
19	9990	9852	8652	8636	8549.6	8511.0	8462.3	8445.7	8436.1
20	12646	11897	9846	9811	9760.6	9710.0	9690.9	9681.4	9675.8
21	11496	11101	9383	9383	9249.9	9219.7	9221.6	9217.8	9216.1
22	9534	9488	8371	8370	8300.8	8281.4	8194.5	8146.7	8122.3
23	8667	8077	7572	7572	7463.7	7440.1	7403.1	7389.6	7378.4
24	8517	7498	7254	7262	7225.8	7195.7	7070.8	6991.3	6973.5
Total	232053	217655	192309	192250	190512.3	189947.5	188154.4	186802.7	186381.6

*4.1. Case 1: All Sources Included.* In the first case, the proposed IMO was utilized in the elucidation of the CEED problem in the microgrid and the case considered both wind and solar energy-powered generators. The optimization results were obtained using the proposed IMA and contrasted with the results obtained from other optimization algorithms. The generation cost calculated by the proposed IMA and other published methods is shown in Table 1 for comparative purposes. Figure 2 shows the convergence characteristics required to mitigate the total generation cost incurred from MA and IMA algorithms. As shown in Figure 2, when cross-verifying the proposed algorithm's cost convergence characteristic, a quicker and more smooth transition was obtained than other optimization techniques considered. Further, Figure 3 shows the comparison results of total generation cost saving from CEED problem when using the mayfly algorithm and other such optimization algorithms. The results achieved from the simulation reveal that the proposed IMA is superior to MA and other optimization algorithms. Moreover, the total generation cost obtained using the proposed IMA algorithm was less than other optimization algorithms. In Figure 4, the total generation cost was obtained using IMA with MA and other published algorithms. It is observed from Figure 4 that improved mayfly optimization algorithm enhanced the total generation cost by 19.68%, 14.37%, 3.08%, 3.05%, 2.17%,

1.88%, 0.94%, and 0.23% over RGM, ACO, CSA, ISA, HIS, IAHS, MHS, and MA, respectively.

*4.2. Case 2: Thermal Power Generating Units without Renewable Sources.* In this case, the mayfly optimization algorithm was utilized to resolve the CEED problem in the microgrid. The case took a total of 3 fossil-fuel-powered thermal generation units under consideration. The optimization results achieved by IMA were contrasted with other such algorithm results. Table 2 shows the total generation cost achieved by the proposed IMA and other such optimization algorithms. The cost convergence profile results are shown in Figure 5 for the proposed IMA and other optimization algorithms. From the results, it can be understood that IMA has promptly converged to the optimal outcome. The comparison results of total generation cost savings are shown in Figure 6 for the CEED problem when using IMA. It shows that IMA achieved better results than MA and other optimization algorithms. The comparison of total generation cost savings obtained using MA and other published algorithms is presented in Figure 7. Figure 7 infers that the total generation cost improved when using the IMA algorithm by 18.52%, 14.67%, 2.93%, 3.3%, 3.27%, 1.32%, and 0.94% over RGM ACO, CSA, ISA, MHS, and MA, respectively. Furthermore, the total generation cost obtained



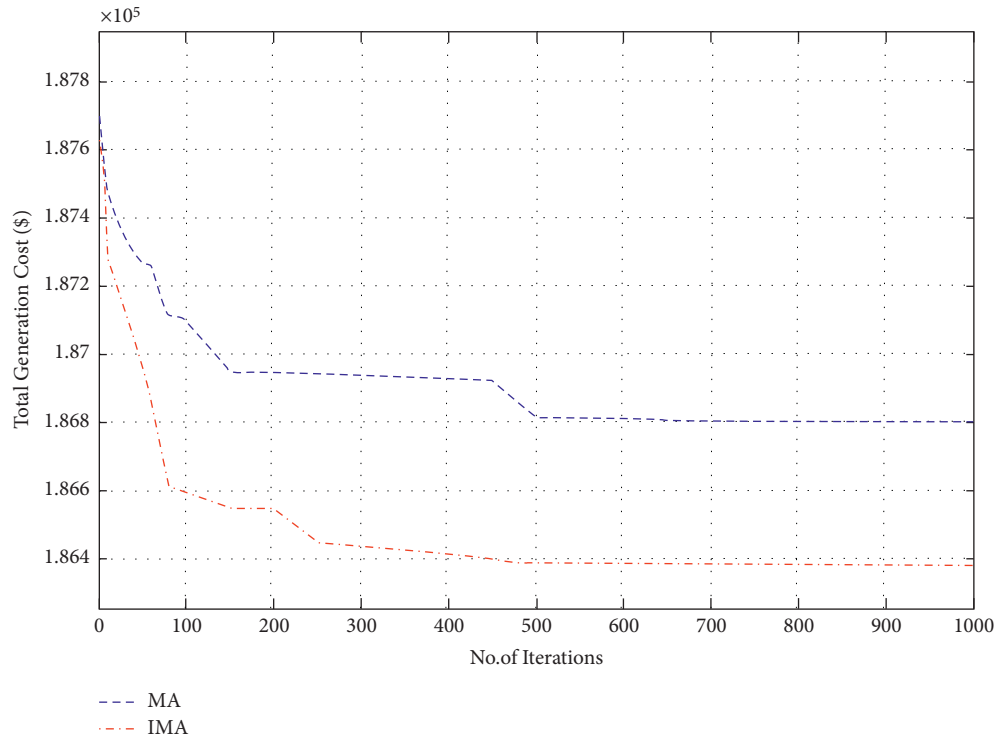


FIGURE 2: Convergence characteristics for total generation cost (case I).

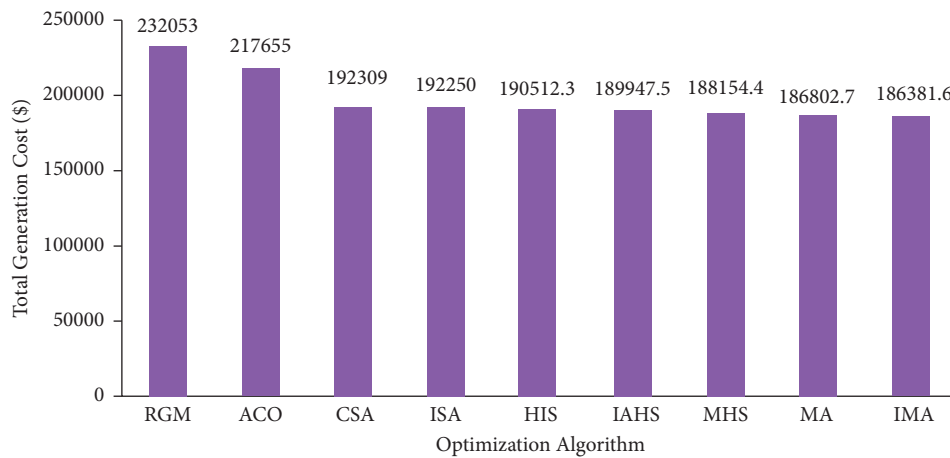


FIGURE 3: Comparison of total generation cost for case 1.

in case 1 can be less than in case 2 due to incorporating renewable energy sources in a microgrid.

4.3. Case 3: Thermal Power Generating Units with Wind Sources Only. Improved mayfly optimization algorithm was deployed in this case to resolve the CEED problem found in microgrids. This case considered fossil-fuel-powered thermal generators in addition to wind sources. IMA and other models (MA, IHS, CSA, and ISA algorithms) were simulated, and the results were compared. Table 3 shows the generation cost calculated for IMA and other such

optimization algorithms. In Figure 8, the author shows the cost convergence characteristic for the optimization algorithms under comparison and the proposed IMA.

Further, Figure 8 also provides an inference; i.e., the convergence characteristic of the proposed LISA strategy II was smooth and quick compared to other strategies. In Figure 9, the researcher compared the total cost saving of IMA and other optimization algorithms from the CEED problem. It is observed from the application results that IMA yielded less total generation cost compared to MA, IHS, CSA, and ISA algorithms. Figure 10 shows the comparison results of operation cost savings obtained using IMA and

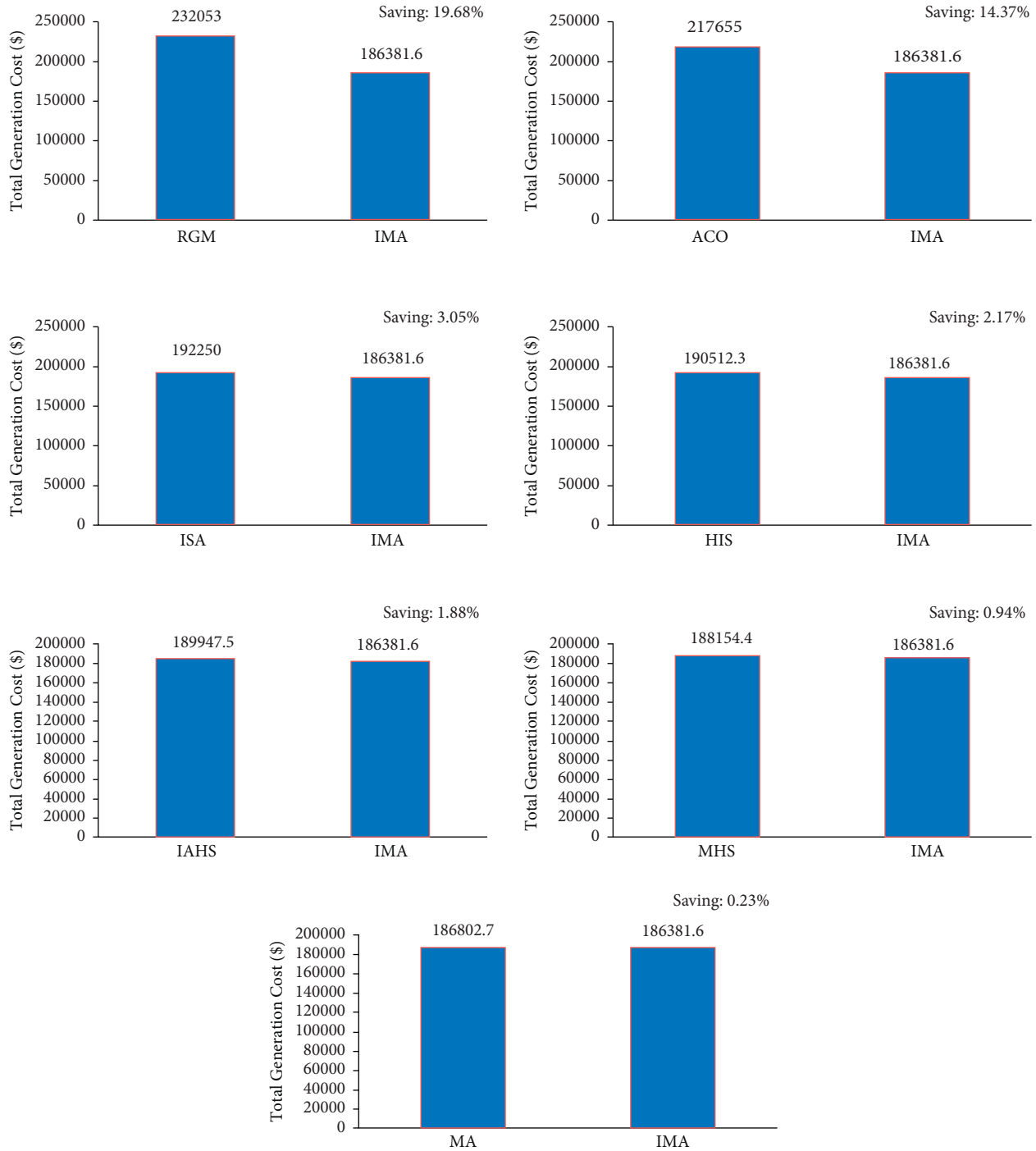


FIGURE 4: Total generation cost saving of CEED problem for case 1.

other published algorithms. It is observed from Figure 10 that IMA improved the operation cost by 3.66%, 3.64%, 1.6%, and 0.55% over CSA, ISA, MHS, and MA, respectively. Further, the total generation cost obtained in this case remains lower than in case 2 because of integrating a wind-powered energy source with the microgrid.

4.4. Case 4: Thermal Power Generating Units with Solar Source Only. The case scenario considered fossil-fuel-powered thermal generating units with solar sources. In this scenario, improved mayfly optimization algorithm was selected to resolve the CEED problem found in microgrids. Simulation results obtained using an improved mayfly optimization

TABLE 2: Optimal generation schedule of microgrid for case 2.

Time (h)	RGM cost (\$/h) [5]	ACO cost (\$/h) [5]	CSA cost (\$/h) [5]	ISA cost (\$/h) [5]	MHS (\$/h) [9]	MA (\$/h)	IMA (\$/h)
1	8490	7317	7179	7179	6977.4	6849.7	6810.2
2	8528	7694	7365	7367	7194.4	7089.6	7061.3
3	8592	7922	7479	7499	7310.3	7223.8	7198.5
4	8675	8117	7598	7608	7429.5	7351.7	7319.3
5	8756	8318	7721	7722	7550.8	7448.2	7416.7
6	8878	8600	7849	7851	7675.4	7567.1	7534.3
7	9005	8768	7978	7978	7802.5	7731.8	7707.4
8	9167	8998	8110	8110	7933.7	7861.8	7829.1
9	10527	10406	8943	8943	8774.3	8697.4	8669.1
10	11867	11347	9540	9540	9380.9	9304.8	9275.4
11	12664	12032	9851	9850	9696.6	9612.5	9590.1
12	13511	12476	10170	10170	10020.0	9973.4	9942.8
13	12664	12032	9850	9746	9696.6	9668.2	9651.6
14	11160	10889	9238	9230	9074.4	8994.6	8971.3
15	10009	9936	8657	8675	8483.5	8401.7	8375.4
16	9167	8998	8110	8109	7933.7	7894.6	7869.2
17	8875	8599	7849	7849	7675.6	7632.8	7605.2
18	9347	9186	8244	8244	8067.5	7992.7	7969.1
19	10009	9936	8657	8657	8483.5	8401.8	8372.4
20	12664	12032	9851	9847	9696.6	9613.8	9589.5
21	11495	11197	9388	9388	9226.1	9148.7	9129.2
22	9540	9479	8377	8379	8203.0	8115.3	7991.5
23	8675	8117	7598	7598	7429.5	7349.3	7317.6
24	8515	7491	7265	7260	7082.8	7005.4	6974.1
Total	240780	229887	202867	202799	198798.4	196930.7	196170.3

algorithm are compared with the outcomes attained by MA and other such algorithms. The total generation cost of the improved mayfly optimization algorithm and other optimization algorithms is presented in Table 4. Figure 11 represents the convergence characteristics obtained to minimize total generation cost using MA and IMA. From Figure 11, it is concluded that the proposed IMA provides steady and quick convergence characteristics. Figure 12 show the comparison results of total generation cost saving achieved by IMA and other optimization algorithms for CEED problem found in microgrids. It is observed from the optimization results that an improved mayfly algorithm provides less total generation cost than other optimization techniques. Figure 13 shows the optimization results of total generation cost saving obtained using IMA and other published metaheuristic optimization algorithms. Figure 13 infers that the proposed IMA algorithm enhanced the total generation cost by 18.5%, 14.7%, 3.44%, 3.41%, 1.43%, and 0.41% over RGM, ACO, CSA, ISA, MHS, and MA, respectively. The authors also conclude that the total generation cost is less in this scenario than in case 2 because of the

incorporation of solar-powered energy sources with the microgrid.

*4.5. Comparison between the Cost Curves of All Scenarios.* Figures 14 and 15 show the comparison results of total generation cost curves under all the scenarios compared to IMA and MA algorithms 24 hours a day. Furthermore, Figure 16 shows the quantitative comparative results of total cost under all the scenarios using IMA. One can notice from Figures 14–16 that case 1 provides a minimum generation cost compared to other scenarios. Also, it can be observed from case 2 that the highest generation cost is obtained in this case. This might be attributed to the reason that renewable energy sources function as negative loads, while the rest are provided by the fossil-fuel-powered thermal generating units only. It reduces the total generation cost. Furthermore, the total generation cost obtained was less in case 3 than in case 4. It could have occurred due to heavy investment costs incurred upon solar power compared to wind power.

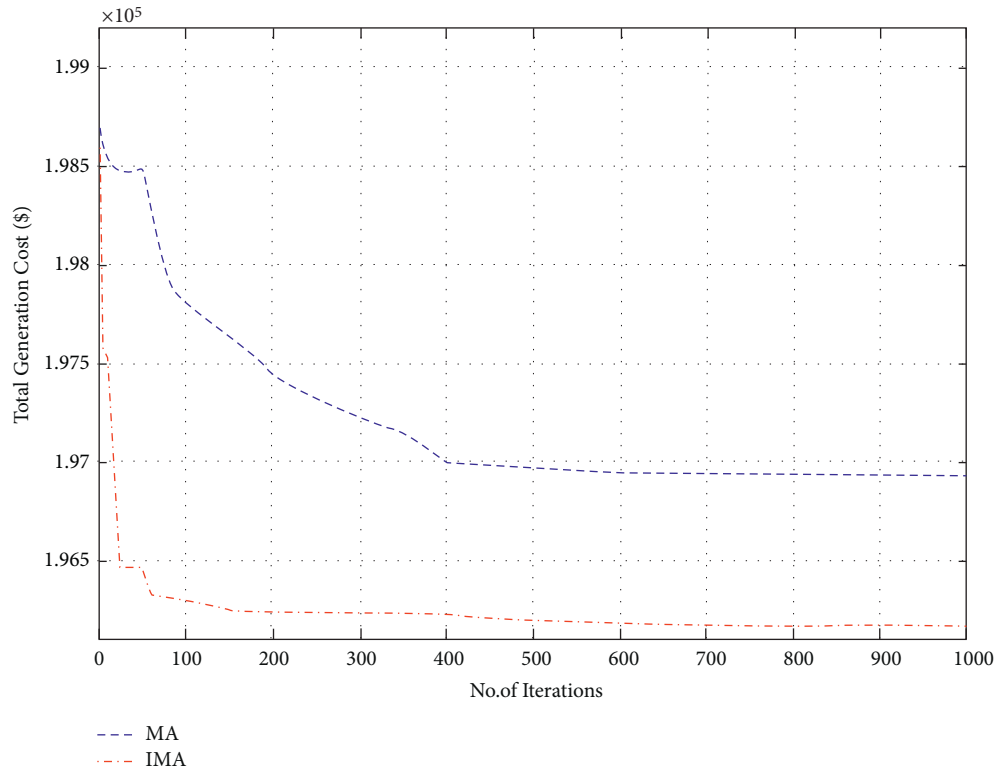


FIGURE 5: Convergence characteristics for total generation cost (case II).

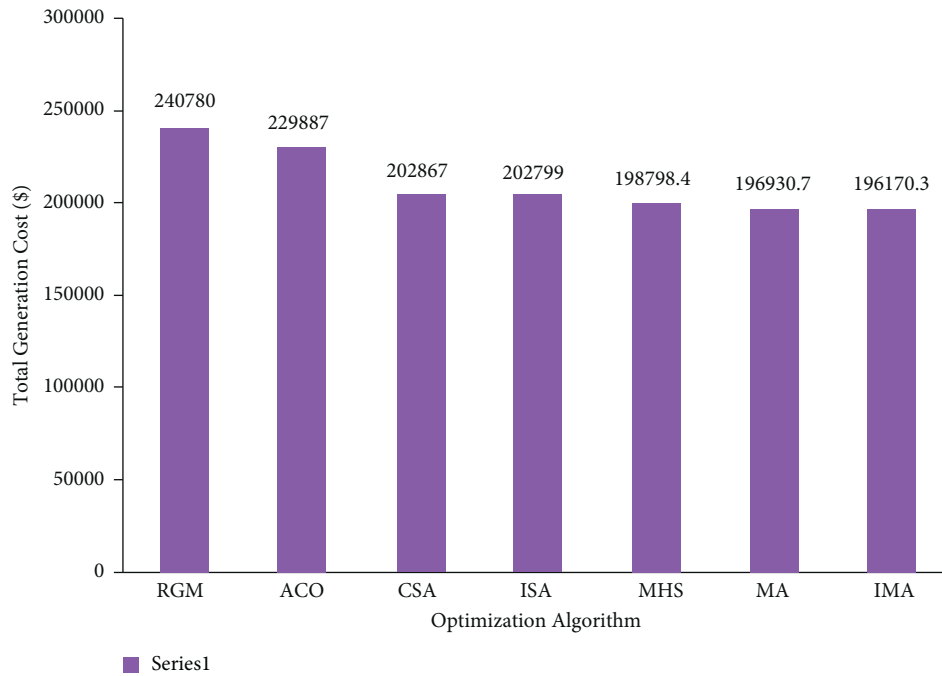


FIGURE 6: Comparison of total generation cost for case 2.

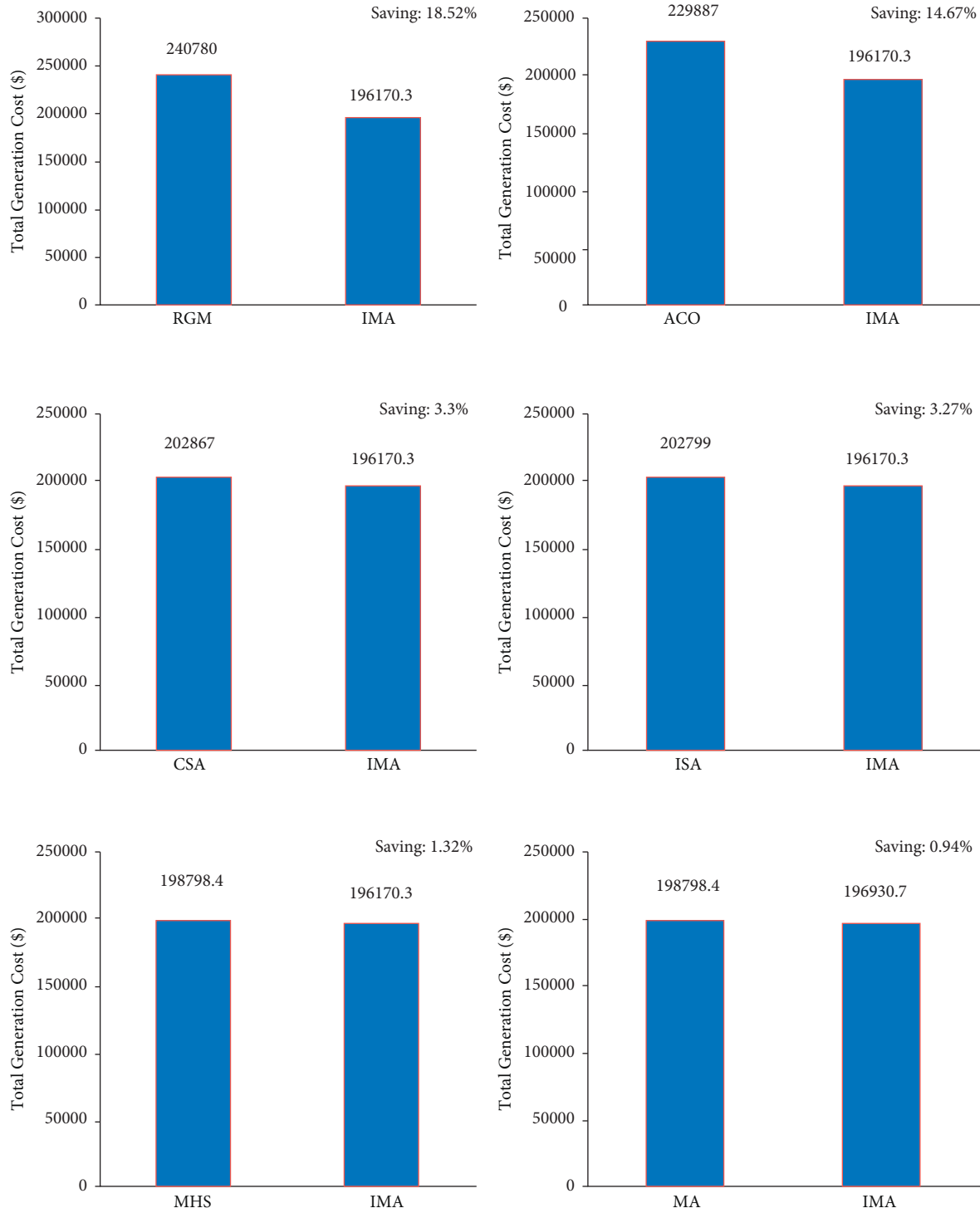


FIGURE 7: Total generation cost saving of CEED problem for case 2.

TABLE 3: Optimal generation schedule of microgrid for case 3.

Time (h)	CSA cost (\$/h) [5]	ISA cost (\$/h) [5]	MHS (\$/h) [9]	MA (\$/h)	IMA (\$/h)
1	7153	7152	6943.4	6851.3	6819.6
2	7203	7199	7010.5	6917.4	6882.7
3	7279	7279	7099.6	7012.6	6990.8
4	7235	7235	7050.2	6972.8	6943.5
5	7544	7545	7377.2	7293.5	7269.3
6	7724	7724	7553.4	7476.8	7435.4
7	7606	7606	7439.4	7355.1	7321.9
8	7443	7443	7278.5	7194.3	7162.8
9	8364	8364	8190.1	8101.6	7784.2
10	9006	9006	8840.7	8753.4	8729.1
11	9454	9461	9295.8	9211.8	9178.6
12	9581	9581	9425.9	9346.9	9313.4
13	9408	9407	9248.7	9168.4	9132.8
14	8933	8933	8766.3	8679.5	8643.7
15	8427	8427	8252.3	8162.9	8123.1
16	7756	7758	7584.2	7495.4	7461.8
17	7761	7761	7590.1	7503.9	7481.6
18	8194	8193	8017.2	7927.1	7893.9
19	8636	8644	8461.9	8361.8	8329.2
20	9845	9842	9690.7	9613.9	9581.3
21	9383	9383	9221.8	9139.6	9107.9
22	8371	8325	8194.7	8127.4	8094.9
23	7572	7571	7403.7	7317.8	7293.5
24	7254	7254	7070.8	6976.7	6943.1
Total	197132	197093	193006.9	190962	189918

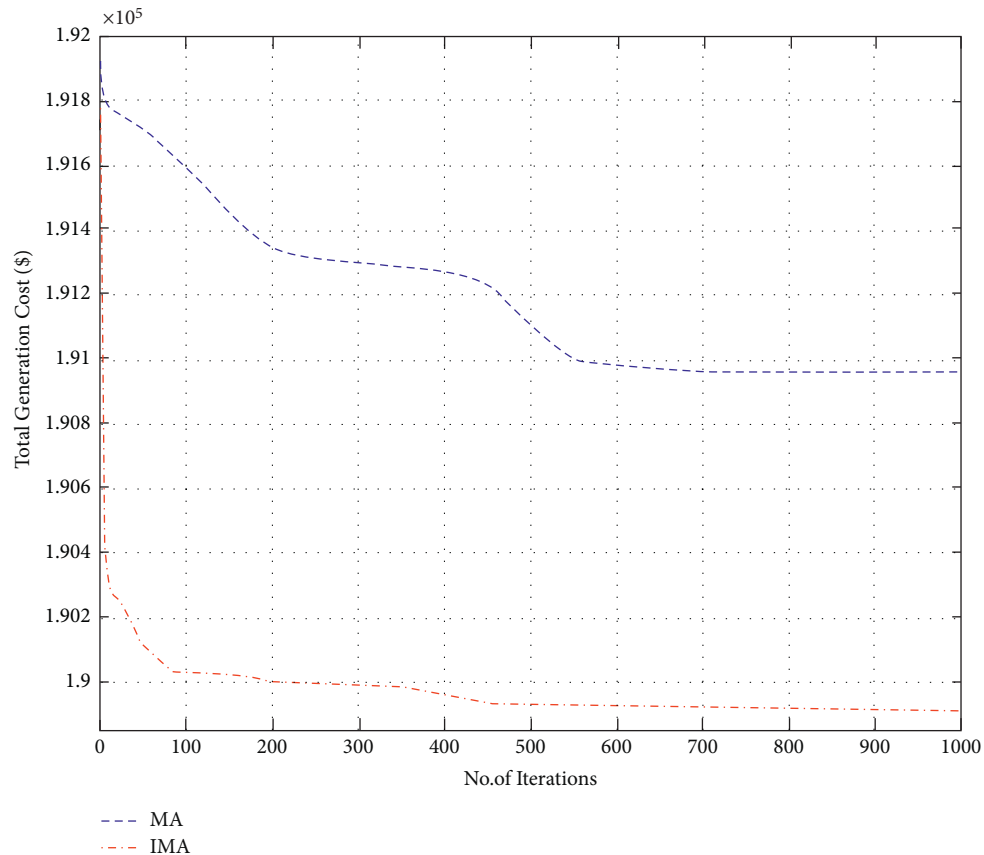


FIGURE 8: Convergence characteristics for total generation cost (case III).

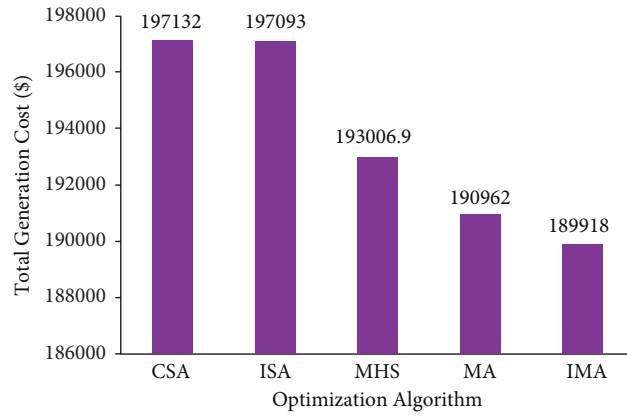


FIGURE 9: Comparison of total generation cost for case 3.

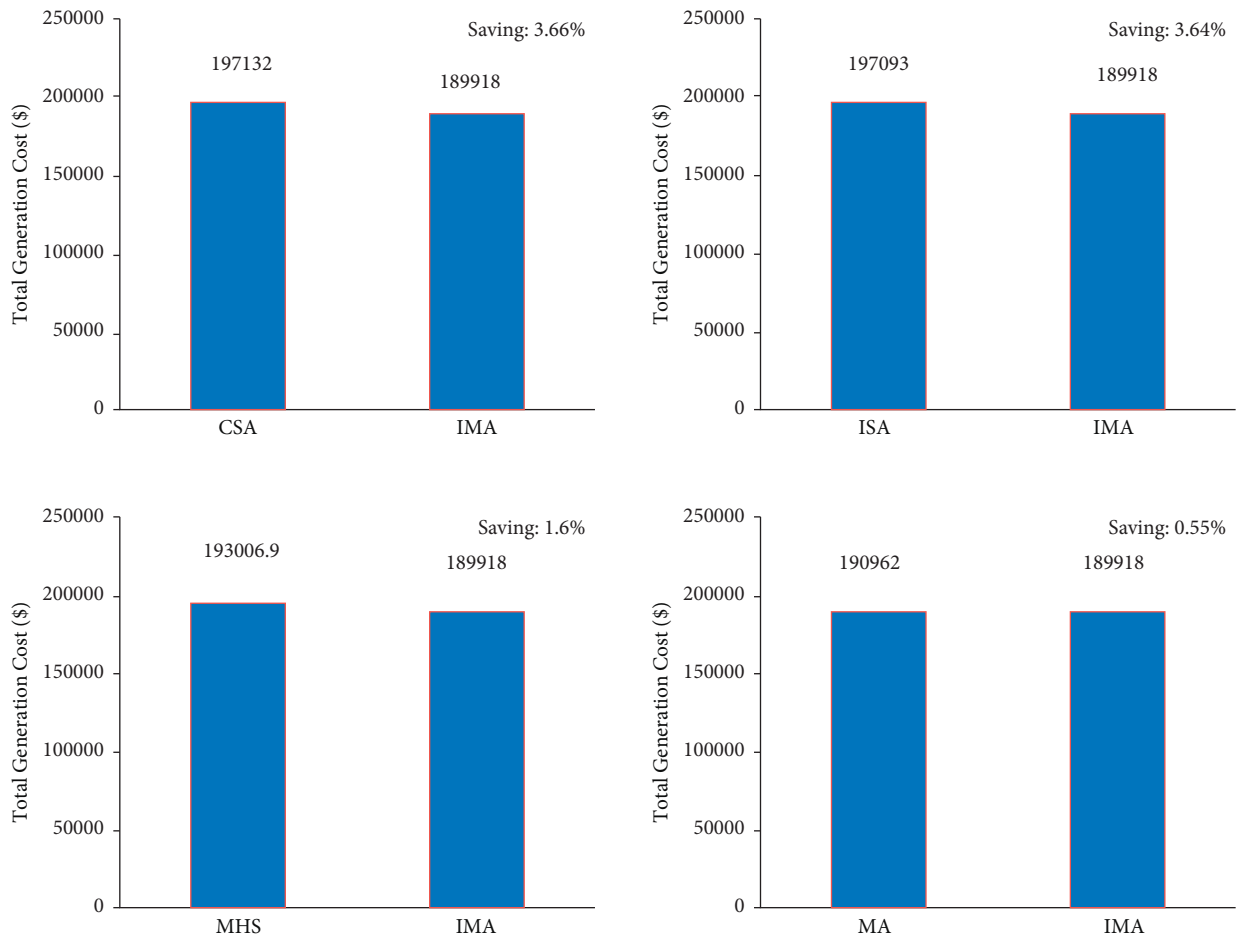


FIGURE 10: Total generation cost saving of CEED problem for case 3.



TABLE 4: Optimal generation schedule of microgrid for case 4.

Time (h)	RGM cost (\$/h) [5]	ACO cost (\$/h) [5]	CSA cost (\$/h) [5]	ISA cost (\$/h) [5]	MHS (\$/h) [9]	MA (\$/h)	IMA (\$/h)
1	8490	7317	7179	7156	6977.4	6885.7	6827.4
2	8528	7694	7365	7364	7194.4	7106.1	7073.2
3	8592	7922	7479	7508	7310.3	7208.9	7174.3
4	8675	8117	7598	7599	7429.5	7335.2	7302.8
5	8756	8318	7721	7721	7550.8	7461.5	7423.9
6	8878	8600	7848	7841	7675.3	7589.8	7554.2
7	8849	8589	7816	7816	7647.2	7559.7	7523.4
8	8969	8559	7692	7692	7530.9	7447.5	7419.2
9	9788	9630	8269	8244	8105.9	8034.4	7994.2
10	10235	10139	8397	8337	8242.2	8187.3	8149.7
11	12153	11648	9620	9634	9465.9	9377.2	9341.8
12	13327	12336	10052	10053	9903.0	9847.2	9829.6
13	10957	10788	8887	8887	8734.9	8660.7	8639.1
14	10153	10012	8467	8467	8305.8	8227.1	8194.7
15	9707	9617	8377	8378	8206.8	8119.4	8094.9
16	9093	8829	7974	7970	7797.9	7707.6	7679.2
17	8810	8279	7608	7608	7444.4	7359.8	7336.2
18	9340	9137	8182	8182	8006.6	7912.4	7884.1
19	10009	9937	8657	8657	8483.5	8391.2	8357.4
20	12664	12032	9851	9849	9696.6	9617.2	9589.7
21	11495	11197	9388	9400	9226.1	9171.2	9143.8
22	9540	9479	8379	8379	8203.0	8112.3	8084.7
23	8675	8117	7598	7596	7429.5	7344.8	7312.2
24	8515	7491	7264	7263	7082.8	6990.4	6943.1
Total	234198	223784	197668	197601	193650.8	191654.6	190872.8

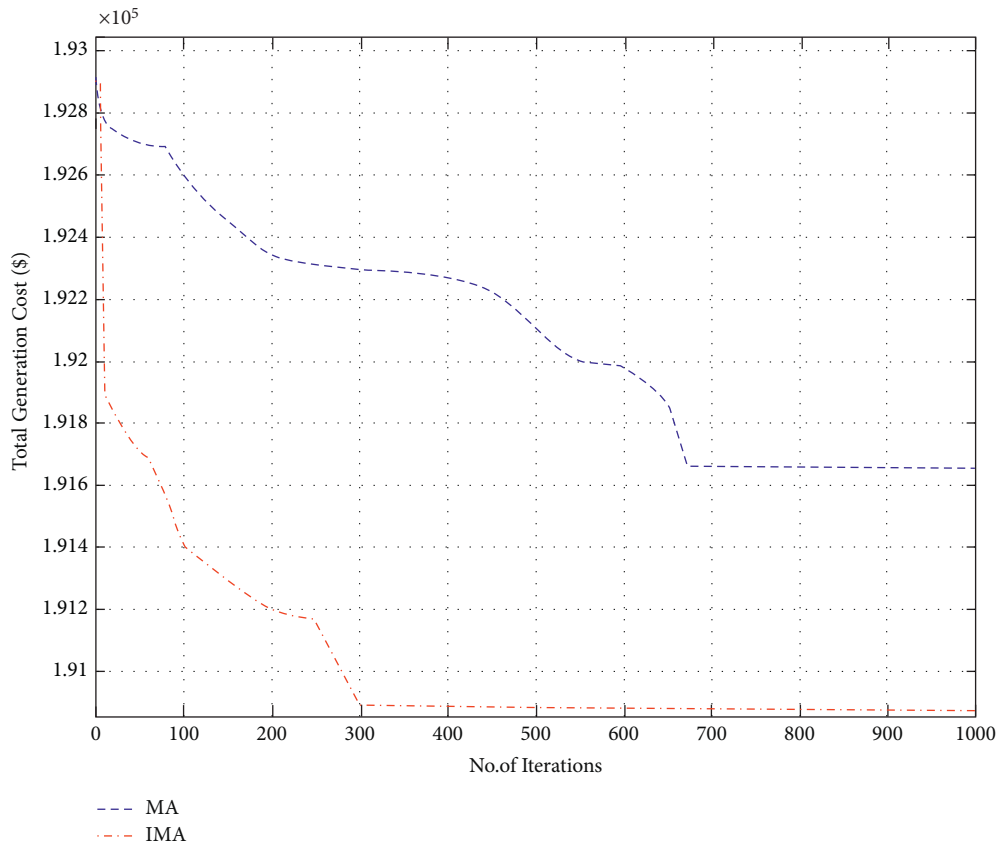


FIGURE 11: Convergence characteristics for total generation cost (case IV).

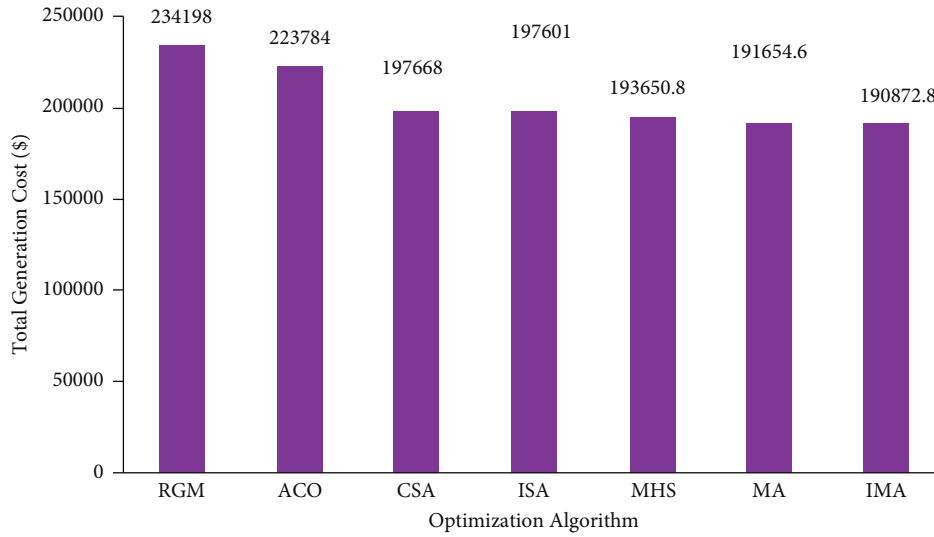


FIGURE 12: Comparison of total generation cost for case 4.

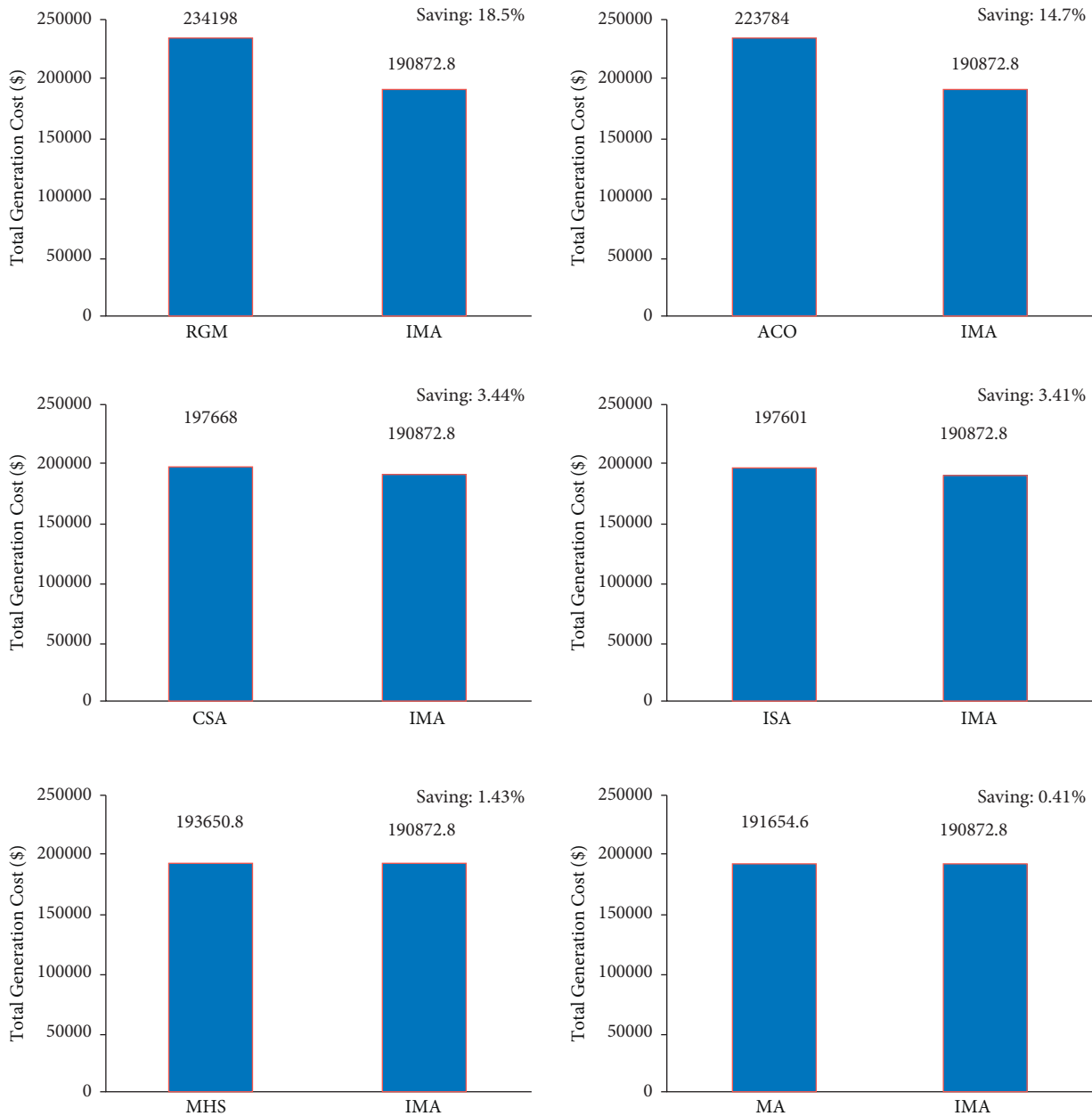


FIGURE 13: Total generation cost saving of CEED problem for case.

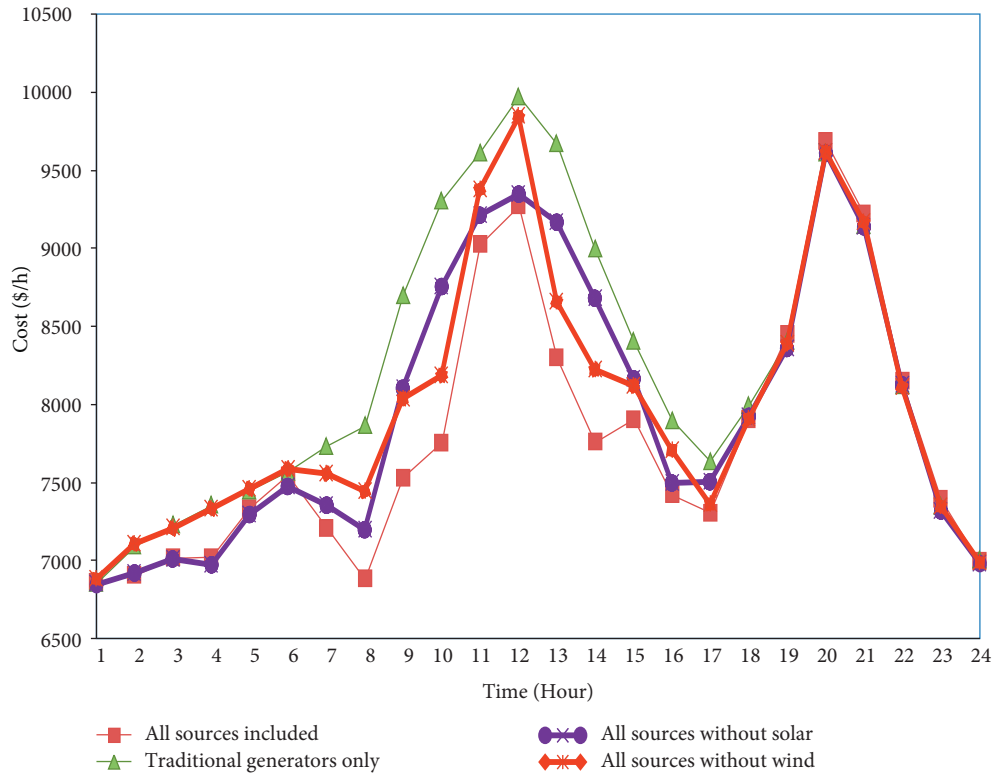


FIGURE 14: Comparison of the cost curve for all cases for 24 hours of a day using mayfly algorithm.

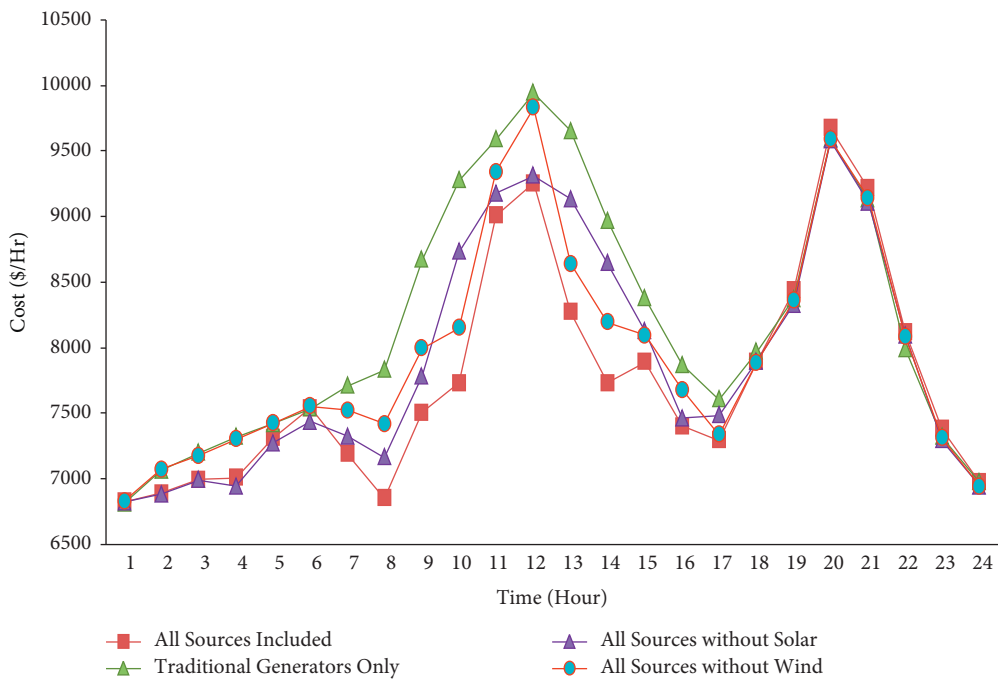


FIGURE 15: Comparison of the cost curve for all cases for 24 hours of a day using improved mayfly algorithm.

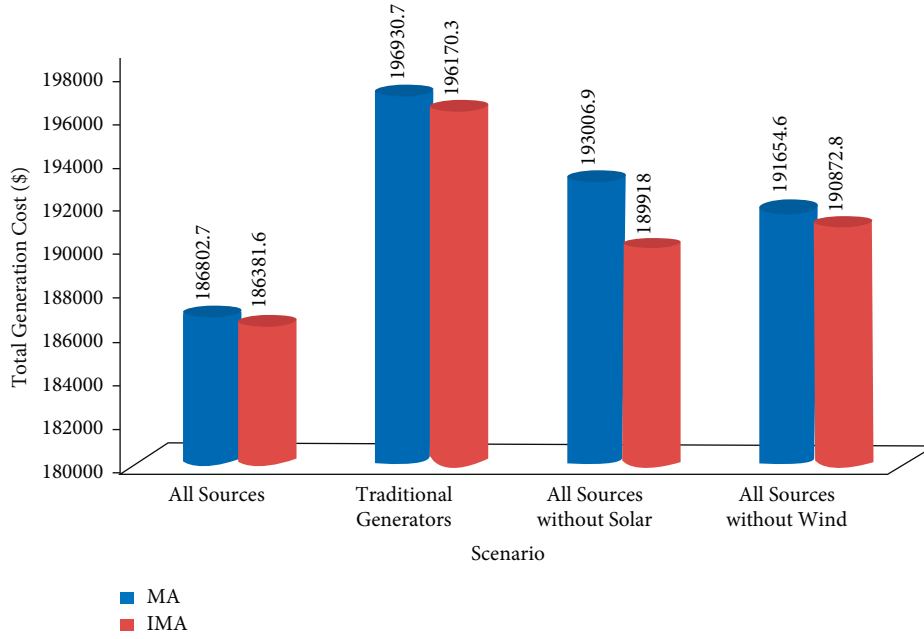


FIGURE 16: Comparison of total generation cost for all cases using MA and IMA algorithm.

## 5. Conclusion

In the current study, the improved mayfly optimization algorithm (IMA) has been implemented to resolve the combined economic emission dispatch (CEED) with renewable energy sources. The study incorporated the proposed IMA as a solution for the CEED problem encountered in the microgrid. Solar and wind power are considered as the cost functions in this study. The proposed IMA algorithm was validated for its supremacy and efficiency in a microgrid model under varying scenarios. The outcomes of IMA and other algorithms were compared and contrasted. The comparison results show that the proposed IMA algorithm is better in cost reduction under all the scenarios. This infers that the proposed IMA is superior, robust, and efficient over other metaheuristic optimization algorithms published earlier. In future, the improved mayfly optimization algorithm can be applied to tackle the CEED problem in grid-connected microgrids comprising battery storage and electric vehicles to accomplish single and multiobjective optimization.

## Nomenclature

### List of Symbols

$a$ :	Annuity coefficient
$C_w$ :	Cost of wind power generating unit
$C_s$ :	Cost of solar-powered generating unit
$d$ :	Coefficient of nuptial dance
$E_i(P_{Gi})$ :	Emission level of the $i^{\text{th}}$ generator
$E_t$ :	Overall emission
$fl$ :	Random walk coefficient
$F_i(P_{Gi})$ :	Fuel cost of the $i^{\text{th}}$ generator

$F_t$ :	Overall fuel cost incurred
$g$ :	Gravity coefficient
gbest:	Global best
iter:	Current iteration
iter <sub>max</sub> :	Maximum no. of iterations
$L$ :	Random number
$l^p$ :	Costs incurred upon investment per unit installed power
$N$ :	Lifetime investment
NG:	No. of generating units
$O^E$ :	Operating and maintenance costs per unit installed power
pbest:	Personal best
$P_{Gi}^{\text{max}}$ :	Maximum output power of generator $i$
$P_{Gi}^{\text{min}}$ :	Minimum output power of generator $i$
$P_D$ :	Total load demand
$P_{Gi}$ , $P_{Gmi}$ , $P_{Gfi}$ :	Power output of $i^{\text{th}}$ generating unit
$P_s$ :	Output power from solar-powered generating unit
$P_w$ :	Output power from wind power generating unit
$r$ :	Interest rate
$r_p$ :	Cartesian distance between $x_i$ and pbest,
$r_g$ :	Cartesian distance between $x_i$ and gbest
$r_{mf}$ :	Cartesian distance between male and female mayflies
TC:	Total operating cost
$\beta$ :	Fixed visibility coefficient
$\Lambda$ :	Price penalty factor
$\Lambda_i$ :	Price penalty factor of $i^{\text{th}}$ generator
$\Delta$ :	Distribution factor
$\delta$ :	Scaling factor
$\Gamma(\cdot)$ :	Gamma distribution function
$\sigma$ :	Standard deviation.

## Abbreviations

BSDE:	Bernstein-search differential evolution
CEED:	Combined economic emission dispatch
ChOA:	Chimp optimization algorithm
CHP:	Combined heat and power
CHPED:	Combined heat and power economic dispatch
CNO:	Collective neurodynamic optimization
DCHPED:	Dynamic combined heat and power economic emission dispatch
DDG:	Dispatchable distributed generator
DE:	Differential evolution
DEED:	Dynamic economic emission dispatch
DSM:	Demand side management
ECED:	Environment constrained economic dispatch
EED:	Economic environmental dispatch
ELD:	Economic load dispatch
EMA:	Exchange market algorithm
FC:	Fuel cell
FCDED:	Fuel constrained dynamic economic dispatch
FP:	Fractional programming
GA:	Genetic algorithm
GWO:	Grey wolf optimizer
HFA/HS:	Hybrid Firefly and Harmony Search
IGWO:	Improved grey wolf optimizer
IMA:	Improved mayfly algorithm
ISA:	Interior search algorithm
LPSP:	Loss of probability of power supply
MA:	Mayfly algorithm
MAACPSO:	Modified adaptive accelerated particle swarm optimization
MBGSA:	Memory-based gravitational search algorithm
MG:	Microgrid
MILP:	Mixed-integer linear programming
MOVCS:	Multiobjective virus colony search
MT:	Microturbine
PPF:	Price penalty factor
PSO:	Particle swarm optimization
PV:	Photovoltaic
QPSO:	Quantum particle swarm optimization
SFS:	Stochastic fractal search algorithm
SGEO:	Social group entropy optimization
SOS:	Sequential optimization strategy
SSER:	System surplus energy rates
WOA:	Whale optimization algorithm
WT:	Wind turbine.

## Data Availability

The data used to support the findings of this study are included in the article. Should further data or information be required, these are available from the corresponding author upon request.

## Disclosure

This study was performed as a part of the Employment of Kombolcha Institute of Technology, Wollo University, Kombolcha, Amhara, Ethiopia.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Acknowledgments

The authors wish to thank the Hindustan Institute of Technology & Science, Chennai, and Vellore Institute of Technology, Chennai Campus, for their support and encouragement to carry out this work.

## References

- [1] N. I. Nwulu and X. Xia, "Optimal dispatch for a microgrid incorporating renewables and demand response," *Renewable Energy*, vol. 101, pp. 16–28, 2017.
- [2] N. Karthik, A. K. Parvathy, R. Arul, and K. Padmanathan, "Economic load dispatch in a microgrid using interior search algorithm," international conference on power and advanced computing, in *Proceedings of the 2019 Innovations in Power and Advanced Computing Technologies (i-PACT)*, Vellore, India, March 2019.
- [3] N. Karthik, A. K. Parvathy, and R. Arul, "A review of optimal operation of microgrids," *International Journal of Electrical and Computer Engineering*, vol. 10, no. 3, pp. 2842–2849, 2020.
- [4] N. Karthik, A. K. Parvathy, R. Arul, and S. Baskar, "A review of optimization techniques applied to solve unit commitment problem in microgrid," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 15, no. 3, pp. 1161–1169, 2019.
- [5] I. N. Trivedi, P. Jangir, M. Bhoje, and N. Jangir, "An economic load dispatch and multiple environmental dispatch problem solution with microgrids using interior search algorithm," *Neural Computing & Applications*, vol. 30, no. 7, pp. 2173–2189, 2018.
- [6] N. Karthik, A. K. Parvathy, and R. Arul, "Multi-objective economic emission dispatch using interior search algorithm," *International Transactions on Electrical Energy Systems*, vol. 29, 2019.
- [7] A. Rajagopalan, P. Kasinathan, K. Nagarajan, V. K. Ramachandaramurthy, V. Sengoden, and S. Alavandar, "Chaotic self-adaptive interior search algorithm to solve combined economic emission dispatch problems with security constraints," *Int Trans Electr Energy Syst*, vol. 29, no. 2, Article ID e12026, 2019.
- [8] N. Karthik, A. K. Parvathy, and R. Arul, "Non-convex economic load dispatch using cuckoo search algorithm," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 5, no. No. 1, pp. 48–57, 2017.
- [9] E. Ehab, "Elattar, Modified harmony search algorithm for combined economic emission dispatch of microgrid incorporating renewable sources," *Energy*, vol. 159, pp. 496–507, 2018.
- [10] B. Dey, B. Bhattacharyya, and F. Pedro García Márquez, "A hybrid optimization-based approach to solve environment constrained economic dispatch problem on microgrid system," *Journal of Cleaner Production*, vol. 307, Article ID 127196, 2021.
- [11] M. Kharrich, O. H. Mohammed, S. Kamel, M. Aljohani, M. Akherraz, and M. I. Mosaad, "Optimal Design of Microgrid Using Chimp Optimization Algorithm," in *Proceedings of the 2021 IEEE International Conference on*

- Automation/XXIV Congress of the Chilean Association of Automatic Control (ICA-ACCA)*, pp. 1–5, Santiago, Chile.
- [12] R. Alayi, M. H. Ahmadi, A. R. Visei, S. Sharma, and A. Najafi, “Technical and environmental analysis of photovoltaic and solar water heater cogeneration system: a case study of Saveh City,” *International Journal of Low Carbon Technologies*, vol. 16, no. 2, pp. 447–453, 2021.
- [13] M. I. Alomoush, “Microgrid dynamic combined power–heat economic-emission dispatch with deferrable loads and price-based energy storage elements and power exchange,” *Sustainable Energy, Grids and Networks*, vol. 26, Article ID 100479, 2021.
- [14] B. Dey, S. K. Roy, and B. Bhattacharyya, “Solving multi-objective economic emission dispatch of a renewable integrated microgrid using latest bio-inspired algorithms,” *Engineering Science and Technology, an International Journal*, vol. 22, no. 1, pp. 55–66, 2019.
- [15] M. I. Alomoush, “Microgrid combined power-heat economic-emission dispatch considering stochastic renewable energy resources, power purchase and emission tax,” *Energy Conversion and Management*, vol. 200, Article ID 112090, 2019.
- [16] T. Wang, X. He, T. Huang, C. Li, and W. Zhang, “Collective neurodynamic optimization for economic emission dispatch problem considering valve point effect in microgrid,” *Neural Networks*, vol. 93, pp. 126–136, 2017.
- [17] F. Moazeni and J. Khazaei, “Dynamic economic dispatch of islanded water-energy microgrids with smart building thermal energy management system,” *Applied Energy*, vol. 276, Article ID 115422, 2020.
- [18] M. Kharrich, S. Kamel, A. S. Alghamdi et al., “Optimal design of an isolated hybrid microgrid for enhanced deployment of renewable energy sources in Saudi arabia,” *Sustainability*, vol. 13, no. 9, p. 4708, 2021.
- [19] A. A. E. Tawfiq, M. O. A. El-Raouf, M. I. Mosaad, A. F. A. Gawad, and M. A. E. Farahat, “Optimal reliability study of grid-connected PV systems using evolutionary computing techniques,” *IEEE Access*, vol. 9, no. 1, pp. 42125–42139, 2021.
- [20] N. B. Roy and D. Das, “Optimal allocation of active and reactive power of dispatchable distributed generators in a droop controlled islanded microgrid considering renewable generation and load demand uncertainties,” *Sustainable Energy, Grids and Networks*, vol. 27, Article ID 100482, 2021.
- [21] X.-G. Zhao, Z.-Q. Zhang, Y.-M Xie, and M. Jin, “Economic-environmental dispatch of microgrid based on improved quantum particle swarm optimization,” *Energy*, vol. 195, Article ID 117014, 2020.
- [22] H. Hou, M. Xue, Y. Xu et al., “Multi-objective economic dispatch of a microgrid considering electric vehicle and transferable load,” *Applied Energy*, vol. 262, Article ID 114489, 2020.
- [23] W. Gil-González, D. Montoya, E. Holguín, A. Garces, and L. F. Grisales-Noreña, “Economic dispatch of energy storage systems in dc microgrids employing a semi-definite programming model,” *Journal of Energy Storage*, vol. 21, pp. 1–8, 2019.
- [24] M. Nemati, M. Braun, and S. Tenbohlen, “Optimization of unit commitment and economic dispatch in microgrids based on genetic algorithm and mixed integer linear programming,” *Applied Energy*, vol. 210, pp. 944–963, 2018.
- [25] P. Pourghasem, F. Sohrabi, M. Abapour, and B. Mohammadi-Ivatloo, “Stochastic multi-objective dynamic dispatch of renewable and CHP-based islanded microgrids,” *Electric Power Systems Research*, vol. 173, pp. 193–201, 2019.
- [26] X. Lu, K. Zhou, and S. Yang, “Multi-objective optimal dispatch of microgrid containing electric vehicles,” *Journal of Cleaner Production*, vol. 165, pp. 1572–1581, 2017.
- [27] R. Alayi, A. Kasaieian, and F. Atabi, “Thermal analysis of parabolic trough concentration photovoltaic/thermal system for using in buildings,” *Environmental Progress & Sustainable Energy*, vol. 38, no. 6, Article ID 13220, 2019.
- [28] R. Alayi, A. Kasaieian, and F. Atabi, “Optical modeling and optimization of parabolic trough concentration photovoltaic/thermal system,” *Environmental Progress & Sustainable Energy*, vol. 39, no. 2, Article ID e13303, 2020.
- [29] M. M. Samy, M. I. Mosaad, and S. Barakat, “Optimal economic study of hybrid PV-wind-fuel cell system integrated to unreliable electric utility using hybrid search optimization technique,” *International Journal of Hydrogen Energy*, vol. 46, no. 20, pp. 11217–11231, 2021.
- [30] F. Nazari-Heris, B. Mohammadi-ivatloo, and D. Nazarpour, “Network constrained economic dispatch of renewable energy and CHP based microgrids,” *International Journal of Electrical Power & Energy Systems*, vol. 110, pp. 144–160, 2019.
- [31] Z. Younes, I. Alhamrouni, S. Mekhilef, and M. Rezasudin, “A memory-based gravitational search algorithm for solving economic dispatch problem in micro-grid,” *Ain Shams Engineering Journal*, vol. 12, no. Issue 2, pp. 1985–1994, 2021.
- [32] Y. Zou, J. Zhao, D. Ding, F. Miao, and B. Sobhani, “Solving dynamic economic and emission dispatch in power system integrated electric vehicle and wind turbine using multi-objective virus colony search algorithm,” *Sustainable Cities and Society*, vol. 67, Article ID 102722, 2021.
- [33] R. Alayi, A. Kasaieian, and A. Njafi, “Optimization and evaluation of a wind, solar and fuel cell hybrid system in supplying electricity to a remote district in national grid,” *International Journal of Energy Sector Management*, vol. 14, no. 2, pp. 408–418, 2020.
- [34] M. M. Samy, M. I. Mosaad, M. F. El-Naggar, and S. Barakat, “Reliability support of undependable grid using green energy systems: economic study,” *IEEE Access*, vol. 9, no. 1, pp. 14528–14539, 2021.
- [35] R. Alayi, F. Zishan, M. Mohkam, S. Hoseinzadeh, S. Memon, and D. A. Garcia, “A sustainable energy distribution configuration for microgrids integrated to the national grid using back-to-back converters in a renewable power system,” *Electronics*, vol. 10, no. 15, p. 1826, 2021.
- [36] M. Basu, “Fuel constrained dynamic economic dispatch with demand side management,” *Energy*, vol. 223, Article ID 120068, 2021.
- [37] M. I. Alomoush, “Application of the stochastic fractal search algorithm and compromise programming to combined heat and power economic-emission dispatch,” *Engineering Optimization*, vol. 52, no. 11, pp. 1992–2010, 2019.
- [38] K. Zervoudakis and S. Tsafarakis, “A mayfly optimization algorithm,” *Computers & Industrial Engineering*, vol. 145, Article ID 106559, 2020.
- [39] A. Y. Abdelaziz, E. S. Ali, and S. M. Abd Elazim, “Implementation of flower pollination algorithm for solving economic load dispatch and combined economic emission

- dispatch problems in power systems,” *Energy*, vol. 101, pp. 506–518, 2016.
- [40] X.-S. Yang and S. Deb, “Multiobjective cuckoo search for design optimization,” *Computers & Operations Research*, vol. 40, no. 6, pp. 1616–1624, 2013.
- [41] H. Hakli and H. Uguz, “A novel particle swarm optimization algorithm with Levy flight,” *Applied Soft Computing*, vol. 23, no. 1, pp. 333–345, 2014.
- [42] N. Karthik, A. K. Parvathy, R. Arul, and K. Padmanathan, “Multi-objective optimal power flow using a new heuristic optimization algorithm with the incorporation of renewable energy sources,” *International Journal of Energy and Environmental Engineering*, vol. 12, no. 4, pp. 641–678, 2021.



## Retraction

# Retracted: A Take-Over Performance Evaluation Model for Automated Vehicles from Automated to Manual Driving

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] L. Yan, J. Chen, C. Wen, P. Wan, L. Peng, and X. Yu, "A Take-Over Performance Evaluation Model for Automated Vehicles from Automated to Manual Driving," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3160449, 16 pages, 2022.

## Research Article

# A Take-Over Performance Evaluation Model for Automated Vehicles from Automated to Manual Driving

Lixin Yan <sup>1</sup>, Jiayu Chen <sup>1</sup>, Chengyue Wen <sup>1</sup>, Ping Wan <sup>1</sup>, Liqun Peng <sup>1</sup>,  
and Xujin Yu <sup>2</sup>

<sup>1</sup>School of Transportation Engineering, East China Jiaotong University, Nanchang 330013, Jiangxi, China

<sup>2</sup>Jiangxi Traffic Monitoring Command Center, Nanchang, Jiangxi 330013, China

Correspondence should be addressed to Lixin Yan; [yanlixinits@163.com](mailto:yanlixinits@163.com)

Received 30 November 2021; Revised 21 December 2021; Accepted 19 March 2022; Published 15 April 2022

Academic Editor: Daqing Gong

Copyright © 2022 Lixin Yan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The evaluation of take-over performance and take-over safety performance is critical to improving the take-over performance of conditionally automated driving, and few studies have attempted to evaluate take-over safety performance. This study applied a binary logistic model to construct a take-over safety performance evaluation model. A take-over driving simulator was established, and a take-over simulation experiment was carried out. In the experiment, data were collected from 15 participants who took over the vehicle and performed emergency evasive maneuvers while performing non-driving-related task (NDRT). Then, to calibrate the abnormal trajectory, the Kalman filter is adopted to filter the disturbed vehicle positioning data and the belief rule-based (BRB) method is proposed to warn irregular driving behavior. The results revealed that the accident rate of male participants is higher than that of female participants in the three frequency take-over experiment, and the overall driving performance of female participants is higher than that of male participants. Meanwhile, medium and high take-over frequencies have a significant effect on the prevention of vehicle collisions. In the take-over safety performance evaluation model, the minimum time to collision (TTC) of 2.3 s is taken as the boundary between the dangerous group and the safety group, and the model prediction accuracy rate is 87.7%. In sum, this study enriches existing research on the safety performance evaluation of conditionally automated driving take-over and provides important implications for the design of driving simulators and the performance and safety evaluation of human-machine take-over.

## 1. Introduction

In recent years, with the strong promotion of communication technology, computer technology, vehicle sensor development, and vehicle positioning system, the research and development of automated driving technology have been continuously promoted. There has been an all-around development whether it is the research on perception, decision-making, control of automated driving vehicles, or the real vehicle test carried out with enterprises as the core. In February 2020, the “Smart Car Innovation Development Strategy” issued by the National Development and Reform Commission in conjunction with 11 departments [1] pointed out that by 2025, the technological innovation, infrastructure, regulations and standards, product

supervision, and network security system of China’s standard smart cars should be formed, and mass production of conditionally automated vehicles is carried out to realize the application of highly automated vehicles in related scenarios. However, at the current stage, there are still many problems to be solved in automated driving technology in terms of driver acceptance and driving safety.

A study on the acceptance of automated vehicles in Australia showed that although most respondents agree with the potential benefits of automated vehicles, they still have considerable concerns about automated vehicles [2]. Yuen et al. [3] used the innovation diffusion theory of perceived value and trust to establish a theoretical model, which identifies potential factors and tests their interrelationships, using a structural equation model to analyze the data obtained from the

questionnaire. The research results point out that the impact of innovation diffusion attributes on public acceptance is completely mediated by the public's perceived value of automated vehicles, and the impact of perceived value on public acceptance is regulated to a certain extent by the public's trust in automated vehicles. By extending the technology acceptance model (TAM) with social and personal factors, Zhang et al. [4] proposed the automated driving acceptance model, which aims to investigate the role of social and personal factors in the acceptance of automated vehicles. The results show that perception factors have a significant impact on user intentions in the initial stage of autonomous driving commercialization, and social influence and initial trust have the greatest impact on user acceptance. Haghzare et al. [5] showed that the older the elderly, the lower the acceptance, but the overall acceptance of automated driving for the elderly is very high.

Regarding the driving safety of automated vehicles, a research report issued by the Insurance Institute for Highway Safety (IIHS) in 2020 pointed out that in the analysis of the causes of the 5,000 major car accidents that occur in the United States each year, the automated vehicles can only reduce the accidents by 1/3. However, according to Tesla's safety report for the first quarter of 2021, Autopilot makes the car nearly nine times safer. After Autopilot turns on, an accident occurs every 6.74 million kilometers, which is 8.66 times lower than the national vehicle statistics. Wang et al. [6] evaluated the safety and effectiveness of 9 common and important automated driving technologies through meta-analysis and tested these technologies in 6 countries to conduct a comprehensive and quantitative assessment of the safety and effectiveness of automated driving technologies. The results show that if all the technologies were implemented in these six countries, an average of 3.4 million accidents can be reduced, of which India has the largest reduction (54.24%). Xiao et al. [7] studied the US cases based on meta-analysis, and it is estimated that intelligent and connected vehicles will reduce the number of fatal accidents by 5% and 13% in 2025 and 2035, respectively.

Currently, automated vehicles are in the transitional stage from L2 to L3. The main factors affecting the take-over performance of L3 automated vehicles are take-over requests (TOR), driver age, and NDRT. Yoon et al. [8] designed 7 types of automated vehicle take-over prompts with visual, auditory, tactile, and mixed sensations. The experimental results show that the take-over effect of driver is poor when there are only visual prompts, and the mixed auditory prompting method allows the driver to take over better. Chu et al. [9] pointed out that among the various possible ways to alert the driver about a TOR, vibrotactile alert provides a significant advantage.

In terms of studying the influence of driver age on take-over performance, the findings of Li et al. [10] show that compared with younger people, older people take over vehicles more slowly and unstable. Wu et al. [11] divided the experimenters into three groups according to their ages and performed non-driving task take-over experiments, and the take-over performance of older drivers was lower than that of younger drivers. However, both the elderly and the young have shown positive views on L3 automation, and the elderly are significantly more active than the young.

In terms of studying the impact of NDRT on take-over performance, Klingegrud et al. [12] designed a human-computer interaction platform to study the ability of drivers to perform NDRT in L4 driving automated vehicles in real traffic environments. NDRT is designed to have visual and cognitive requirements, and manual interaction is required. The results show that drivers can participate in NDRT to a large extent. Rauffet et al. [13] carried out a take-over experiment in a driving simulator where the non-driving task was playing video games. The ratio of browsing time in the game and the time between game sessions were used as indicators of game participation. The survey results showed that the participants were highly involved in NDRT, and the average take-over time was longer than if they were not engaged in NDRT. Ou et al. [14] designed experiments to allow drivers who are immersed in NDRT to detect TOR and quickly brake. The research results show that advanced predictive interfaces that provide directional information can significantly improve take-over performance.

When studying the take-over test of human-machine codriving, it is necessary to use a simulation platform to build relevant driving scenarios. Zhang and Zhu [15] used a longitudinal research design on a driving simulator based on Unity's complete cab to study the changes in driver state and behavior during multiple sessions of automated vehicle operation. Calvi et al. [16] use STISIM for driving scene simulation, plus a driving controller (wheels, pedals, and gears) connected to the workstation of the control system to convert it into a driving simulator. Experiments are conducted to study the behavior of the driver after the driver is inattentive and participating in secondary tasks during highly automated driving. Yoon and Ji's [17] driving simulator is based on the City Car software simulation scene, equipped with a real car seat, a Logitech racing force feedback wheel and pedal, and a 55-inch Samsung smart TV. In addition, an SMI eye tracker was used to collect the eye movement data of the participants.

In the research on the take-over performance of automated driving, Zhou et al. [18] proposed using eye tracking and self-report data to predict the situation perception during the transition of conditionally automated driving. The tree-integrated machine learning model light gradient boosting machine (LightGBM) is used to predict the situation perception. Then, the Shapley additive explanation (SHAP) value of the individual predictor variables in the model is calculated. The research on automated driving take-over performance is mainly based on objective data and selected relevant indicators for evaluation. In the study of Wiedemann et al. [19], the performance evaluation of the driver is divided into two parts: horizontal control and vertical control. The horizontal control index selects the vehicle lateral displacement and steering wheel angle, and the longitudinal control index selects the vehicle velocity. Du et al. [20] predicted the take-over performance of drivers before the TOR by analyzing physiological data such as driver's heart rate index, skin electrical response index, and eye-tracking index, and external environmental data such as scenario type, traffic density, and lead time of TOR. The study found that the random forest classifier can better

predict the driver's take-over behavior, using 3 s as the optimal time window for predicting take-over performance, with an accuracy rate of 84.3%.

In the human-machine codriving take-over experiment, the importance of take-over safety evaluation analysis cannot be ignored. Lin et al. [21] used binary logistic regression to establish a take-over safety evaluation model to evaluate the safety of L3 automated vehicles. The research results show that under the condition that the TOR time is 7 s, the main factors affecting the safety of take-over are the take-over response time and second tasks. The established take-over safety evaluation model has a prediction accuracy of 85.5%. As domestic automated driving is at an immature stage, there are relatively few studies on the safety evaluation of take-over.

Through the above research, it is found that the selection of driver behavior indicators in human-machine codriving is mainly the driver's visual characteristics and behavior characteristics. At the same time, the driving simulation platform can effectively reduce the cost of the experiment and ensure the safety of the experiment. Therefore, this article uses a driving simulator designed based on the CARLA driving simulation platform and Logitech G29 force feedback steering wheel pedal set to study the driver's take-over performance and its influencing factors when the driver performs the take-over operation under the L3 of automated driving (conditionally automated driving). Comparing the conditions of the same non-driving-related tasks and different take-over frequencies, the evaluation of the driver's take-over performance is completed by analyzing factors such as the average time to complete the tasks, the minimum TTC, the distance to obstacle, and the maximum braking acceleration. Combining the gender, age, temperament type, driving style, and other conditions of driver, this study studies the commonality and characteristics of automated driving take-over performance of different groups of people. Finally, the binary logistic model is applied to evaluate the safety of the test and find out the factors that affect the safety of the take-over.

## 2. Methodologies

### 2.1. Data Preprocessing Method

**2.1.1. Kalman Filter Method.** The Kalman filter is an efficient autoregressive filter, which can predict the dynamic state of the system in a series of incomplete and noisy measurement values. The Kalman filter can estimate the unknowns from the measured values at different times regarding their joint distribution, so the result obtained is more accurate than the prediction method based on only a single measured value. The Kalman filter is based on the state matrix, the localization data are the input of the filter, the system state prediction data are the output, and the prediction equation and the measurement equation are used to establish the relationship between input and output, which is a method of calculating the value of the system state with input and

output. The Kalman filter is composed of three parts: state prediction equation, state observation equation, and recursive equation. From  $k - 1$  to  $k$  moments, the calculation process is as follows.

The state prediction equation is as follows:

$$X_k = AX_{k-1} + B\mu_k + w_k, \quad (1)$$

where  $X_k$  represents the state vector of the system at time  $k$ ,  $A$  is the state transition matrix,  $B$  is the input gain matrix,  $w_k$  is the mean value of 0, and  $Q$  is the covariance matrix, and equation obeys the process noise of normal distribution.

The state observation equation is as follows:

$$Z_k = HX_k + v_k, \quad (2)$$

where  $Z_k$  represents the observation vector of the system at time  $k$ ,  $H$  is the measurement matrix,  $v_k$  is the mean value of 0, and  $R$  is the covariance matrix, and equation obeys the measurement noise of normal distribution.

The Kalman filter has three common kinematic models: constant velocity (CV) model, constant turn rate and acceleration model (CTRA), and constant acceleration (CA) model. Compared with the CA model, the calculation accuracy of CTRA model is improved slightly, but the calculation amount is increased significantly. Considering the calculation efficiency and accuracy comprehensively, the CA model is selected as the kinematic model of the Kalman filter in this study. Combined with the state prediction equation,  $X_k$  can be expressed as follows:

$$X_k = AX_{k-1} + w_k, \quad (3)$$

$X_k = [x_t, x'_t, x''_t, y_t, y'_t, y''_t]^T$ , where  $x_t, x'_t, x''_t$  represents the position, velocity, and acceleration of the system in the  $x$ -direction at time  $t$ , and  $y_t, y'_t, y''_t$  represents the position, velocity, and acceleration in the  $y$ -direction.

The recursive equation is as follows:

The linear estimate of the system at time  $k$  is predicted using the state value at time  $k - 1$ :

$$\hat{x}' = A\hat{x}_{k-1} + B\mu_k. \quad (4)$$

The new variance predicted by the error covariance and system noise  $Q$  at the last moment is as follows:

$$P'_k = AP_{k-1}A^T + Q. \quad (5)$$

The state correction process is as follows:

$$K_k = P'_k H^T (HP'_k H^T + R)^{-1}, \quad (6)$$

$$\hat{x}_k = \hat{x}'_k + K_k (Z_k - H\hat{x}'_k).$$

The state covariance estimates are updated:

$$P_k = (1 - K_k H)P'_k, \quad (7)$$

where  $\hat{x}_k$  is the Kalman estimation value,  $\hat{x}'$  is the prediction value,  $P_k$  is the Kalman estimation error covariance matrix,  $P'_k$  is the prediction error covariance matrix, and  $K_k$  is the Kalman gain.

**2.1.2. The Belief Rule-Based (BRB) Method.** The general expression form of the BRB method is as follows: if  $A_1^k \wedge A_2^k \wedge \dots \wedge A_T^k$ , then  $(C_1, \beta_{k1}), (C_2, \beta_{k2}), \dots, (C_N, \beta_{kN})$  with a rule weight  $\theta_k$  and attribute weights  $\delta_1, \dots, \delta_T$ ,  $\sum_{l=1}^N \beta_{kl} \leq 1$ , where  $A_1, \dots, A_T$  represents the premise attribute of the BRB,  $C_1, \dots, C_N$  represents the evaluation level of the evaluation result, and  $A_j^k \in \{C_{j1}, \dots, C_{jN}\}$  is an evaluation grade relative to the premise attribute  $A_j$ ,  $j \in \{1, \dots, N\}$ .  $\beta_{kl}$  is the confidence level of the result under the  $k$ th rule in the evaluation level  $C_l$ ,  $l \in \{1, \dots, N\}$ , and  $T_l$  is the total number of premise attributes under the  $k$ th rule. If  $\sum_{l=1}^N \beta_{kl} = 1$ , then the BRB is considered complete, otherwise it is incomplete. In this case, to establish a BRB, discrete evaluation levels such as high, medium, and low need to be defined for each premise attribute and result. Several types of parameters in the BRB system structures such as confidence, rule weights, and attribute weights can be obtained by training the collected sample data (velocity and acceleration). The sample data are shown in Table 1. The trained BRB system can now be used to predict the authenticity of the target.

**(1) BRB Input Conversion.** The input conversion of a premise attribute value refers to converting this value into different confidence levels and assigning these confidence levels to different reference values of the premise attribute. This is equivalent to converting an input value into a confidence distribution corresponding to the reference value of the premise attribute. In particular, the input value of a premise attribute  $P_i$  (its confidence level is set to  $x_i$ ) can be transformed by the following confidence distribution:

$$S(P_i, x_i) = \{(h_{in}, \alpha_{in}), n = 1, \dots, n_i\}, \quad i = 1, \dots, T_i, \quad (8)$$

where  $S$  represents the distribution of the estimated confidence level assigned to the input value of the premise attribute,  $h_{in}$  (the  $i$ th value) is the  $n$ -th reference value of the input premise attribute  $P_i$ , and  $\alpha_{in}$  ( $\alpha_{in} \geq 0$ ) is the confidence level corresponding to the reference value  $h_{in}$ . Among them,  $\sum_{n=1}^{n_i} \alpha_{in} \leq 1$  ( $i = 1, \dots, T_i$ ), where  $n_i$  is the quantity of the reference value.

In this study, the velocity and acceleration of the target point are selected as the premise attributes in the BRB framework. Their input values can be obtained from the radar data of the simulator, and the input values are converted into the membership degree of the corresponding premise attribute reference value. The reference values corresponding to velocity and acceleration are velocity (fast, normal, and low) and acceleration (high, medium, and low). At this point, these linguistic values (evaluation levels) can be assigned to the confidence level  $\alpha_{in}$  through expert evaluation, which is then distributed over the different reference values of the premise attributes  $h_{in}$  (high (H), medium (M), and low (L)) in terms of confidence level  $\alpha_{in}$ . The input conversion process of the above can be explained by formula.

High  $\geq x_i \geq$  Medium,

$$\begin{aligned} \text{Medium} &= \frac{\text{High} - x_i}{\text{High} - \text{Medium}}, \text{High} \\ &= (1 - \text{Medium}), \text{Low} = 0, \end{aligned} \quad (9)$$

$$\text{Medium} > x_i \geq \text{Low}, \text{Low} = \frac{\text{Medium} - x_i}{\text{Medium} - \text{Low}},$$

$$\text{Medium} = (1 - \text{Low}), \text{High} = 0.$$

Before the input conversion, the reference point value of the premise attribute needs to be set. Assume that the reference value distribution of the two premise attributes is as follows: velocity = {(fast, 41.7), (normal, 27.8), (low, 16.7)}; acceleration = {(high, 8), (medium, 5), (low, 3)}. Then, formula (9) is combined to complete the calculation of the confidence level  $\alpha_{in}$ . For example, a sample with velocity = 36.34 m/s and acceleration = 4.73 m/s<sup>2</sup> is selected, and then, the degree of velocity belonging to (fast, 41.7), (normal, 27.8), (low, 16.7) is (0.614, 0.386, 0). That is, medium = (41.7 - 36.34)/(41.7 - 27.8) = 0.386, high = 1 - 0.386 = 0.614, and low = 0. In the same way, the degree of acceleration belonging to (high, 8), (medium, 5), (low, 3) is (0, 0.865, 0.135).

**(2) Calculation of Activation Weight.** Generally, in the BRB, the connecting symbol “ $\wedge$ ” is usually applied to represent the logical relationship of the premise attribute. It means that only when all the premises in the rule are activated the results obtained at this time can be credible. According to the calculation of the above confidence distribution, the activation weight  $\omega_k$  under the  $k$ th rule can be calculated by the formula:

$$\begin{aligned} \omega_k &= \frac{\theta_k \prod_{i=1}^{T_k} (\alpha_{ik})^{\bar{\delta}_i}}{\sum_{l=1}^L \left[ \theta_l \prod_{i=1}^{T_l} (\alpha_{il})^{\bar{\delta}_i} \right]} \\ \bar{\delta}_i &= \frac{\delta_i}{\max_{i=1, \dots, T_k} \{\delta_i\}}. \end{aligned} \quad (10)$$

Among them,  $\alpha_{ik}$  ( $i = 1, \dots, T_k$ ) is the input confidence level of the  $i$ th premise attribute, representing the individual matching degree and evaluating the reference value  $A_i^k$  under the  $k$ th rule.  $\theta_k \prod_{i=1}^{T_k} (\alpha_{ik})^{\bar{\delta}_i}$  is the joint matching degree, which reflects the matching degree between the input value of the entire premise attribute and the reference value under the  $k$ th rule.  $\bar{\delta}_i$  (or  $\delta_i$ ) represents the attribute weight. It is worth noting that if  $\bar{\delta}_i = 0$ , then  $(\alpha_{ik})^{\bar{\delta}_i} = 1$ , which means that a premise attribute of zero importance will not have any effect on the activation weight; if  $\bar{\delta}_i = 1$ , then  $(\alpha_{ik})^{\bar{\delta}_i} = \alpha_{ik}^k$ , which means most of the premise attributes have a significant impact on the activation weight. With the calculation of the activation weights completed, the estimated output for a specific input vector can be inferred using ER theory.

TABLE 1: Experimental data sample.

Sample	1	2	3	4	5	6	7	8	9	10
Velocity (m/s)	36.21	36.34	36.63	37.04	37.59	38.48	39.44	40.66	41.88	43.13
Acceleration (m/s <sup>2</sup> )	4.86	4.73	4.88	4.61	4.69	4.77	4.94	4.83	4.66	4.72

(3) *ER Inference Output*. The ER theory is applied to collect all the premise attribute data groups under the  $L$  rule, to obtain the confidence level of each reference value in the result attribute through the given premise attribute  $P_i$  input value. This study adopts the analytical ER algorithm, and its output result is composed of the reference value of the result attribute, as shown in the following formula:

$$\beta_j = \frac{\prod_{k=1}^L (\omega_k \beta_{jk} + 1 - \omega_k \sum_{j=1}^N \beta_{jk}) - \prod_{k=1}^L (1 - \omega_k \sum_{j=1}^N \beta_{jk})}{\sum_{j=1}^N \prod_{k=1}^L (\omega_k \beta_{jk} + 1 - \omega_k \sum_{j=1}^N \beta_{jk}) - (N-1) \prod_{k=1}^L (1 - \omega_k \sum_{j=1}^N \beta_{jk}) - \prod_{k=1}^L (1 - \omega_k)} \quad (12)$$

The final merged result or the output expression obtained by ER inference is  $\{(C_1, \beta_1), \dots, (C_N, \beta_N)\}$ , where  $\beta_j$  is the final confidence level that belongs to the  $j$ th reference value  $C_j$  in the result attribute. The output distributions selected in this study are the real target and the false target, respectively. According to the sample data in Table 1, combined with formula (12), the final BRB after training is shown in Table 2.

The BRB in Table 2 describes the causal relationship between velocity and acceleration and the true target position. For example, rule 3 represents that the value of velocity is in the fast range and the value of acceleration is in the medium range at the current moment. The confidence level of the real target position and the false target position is 93.65% and 6.35%, and the confidence of this rule is 1. Therefore, the results obtained under this rule can be fully believed. The rest of the rules can be explained in the same way. Consequently, it has a certain degree of reliability to combine velocity and acceleration to test the trajectory of the target vehicle.

**2.2. Binary Logistic Regression Model.** In Natural Inheritance published in 1889, Francis Galton, a famous British biologist and statistician, first proposed that “The logistic regression model derived from the logistic curve are probabilistic regression, belonging to generalized linear regression.” The curve of the logistic function is a monotonously increasing function with no breakpoints but good continuity. The horizontal coordinate range (sample input range) of the curve is  $(-\infty, +\infty)$ , and the vertical coordinate ranges  $(0, 1)$ . The function distribution trend is exactly what many probability problems need to be or not. The derived logistic regression function, also known as the growth function, more often uses a binary dependent variable. The relationship between variables is used to make classification judgments on the prediction results.

The maximum-likelihood method is usually adopted for the estimation of logistic regression parameters. The basic

$$O(Y) = S(P_i) = \{(C_j, \beta_j), \quad j = 1, \dots, N\}, \quad (11)$$

where  $\beta_j$  represents the confidence level corresponding to a result reference value  $C_j$ . The calculation of  $\beta_j$  is obtained by the analysis formula (12) of the ER algorithm.

idea of the method is to establish the likelihood function and the logarithmic likelihood function first and then solve the parameter value corresponding to the maximum logarithmic likelihood function. The estimated value obtained is the called maximum-likelihood estimation of the parameter. It can be seen from formula (13) that the logistic model establishes the relationship between the probability of event occurrence and explanatory variables.

$$\ln \frac{p}{1-p} = \alpha + X\beta + \varepsilon, \quad (13)$$

where  $p$  is the probability of the event,  $\alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$  is the

intercept term of the model,  $\beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix}$  is the parameter to

be estimated,  $X = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nk} \end{pmatrix}$  is the explana-

tory variable, and  $\varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$  is the error term.

### 3. Data Collection and Analysis

#### 3.1. Experiment Design and Procedure

**3.1.1. Apparatus.** The driving simulator uses the CARLA driving simulation platform, 3ds Max, and Unreal 4 modeling software to construct virtual driving scenes, external Logitech G29 force feedback steering wheel pedal package, DXRacer car seat, two 40-inch screen 2K monitors, and one set of stereo, as shown in Figure 1. The Logitech G29 force feedback steering wheel pedal package includes a steering wheel, a manual gear, and a racing pedal with a clutch. The



TABLE 2: BRB after training.

Number of rules	Rule weight	Velocity	Acceleration	Conclusion
1	1	Fast	Medium	[0.9164, 0.0836]
2	1	Fast	Medium	[0.9707, 0.0293]
3	1	Fast	Medium	[0.9365, 0.0635]
4	1	Fast	Medium	[0.9251, 0.0749]
5	1	Fast	Medium	[0.9339, 0.0441]
6	1	Fast	Medium	[0.9361, 0.0639]
7	1	Fast	Medium	[0.9228, 0.0772]
8	1	Fast	Medium	[0.9650, 0.0350]
9	1	Fast	Medium	[0.0632, 0.9368]



FIGURE 1: Driving simulator.

steering wheel has a 900-degree steering range, which can simulate real driving behavior to the greatest extent. It also provides programmable keys and direction control keys. Two large-screen high-definition monitors serve as visual feedback devices to display virtual driving scenes. The stereos act as auditory feedback devices and take into account the Doppler effect to simulate real driving and environmental sound.

**3.1.2. Participants.** A total of fifteen participants with a valid driver's license were recruited for the driving simulator experiment. Nine participants were males, and six participants were females. The average age of the participants was 22.07 years (SD = 1.69 years). They had on average 2.27 years of driving experience (SD = 1.34 years). Before the experiment, the participants were required to tell their health conditions such as illness, fatigue, and drug misuse. After the experiment, monetary compensation (100 RMB) was offered for their participation.

The participants were required to take over for 1 time, 5 times, and 9 times, and the interval of each experiment was 3 minutes. The NDRT of the experiment was set to watch a video, which was playing on the tablet computer on the side of the control platform in the vehicle. The TOR time refers to the TTC between the self-vehicle and the vehicle or obstacle on the road ahead, which represents the urgency of the taking over. In this experiment, the TTC is 5 s, leaving enough time for the participants to react.

**3.1.3. Scenario Design.** The test road is the same two-way four-lane urban simulation road with a total length of about 5 km. The limitation of road velocity is 30 km/h, and the automated driving velocity is about 20–24 km/h. The test includes three different take-over frequencies (low take-over

frequency, medium take-over frequency, and high take-over frequency). The duration is 13 minutes, 16 minutes, and 18 minutes, respectively. This article mainly studies dangerous situations, the participant switches the automated driving mode to manual driving mode, takes over the vehicle, and bypasses obstacles or brakes. In this study, a simplified “ghost probe” driving scene was designed in the simulator. When the vehicle is driving to a certain position in the automated driving mode, a pedestrian crossing the road suddenly appears on the roadside ahead of the road. At this time, the vehicle take-over system will issue a TOR. The participant needs to take over and break the vehicle; otherwise, the vehicle will collide with a pedestrian in front of the vehicle and cause a traffic accident as shown in Figure 2.

$x_1$  is the lateral position where pedestrian appear,  $x_2$  is the lateral position of the vehicle head, and the vehicle is moving at a constant velocity. The upper and lower limits of the vehicle velocity  $v_{\diamond}$  are set as  $v_1$  and  $v_2$ , respectively,  $y$  is the longitudinal distance between the pedestrian and the vehicle, and  $l$  is the width of the vehicle. Pedestrians cross the road vertically at velocity  $v$ , and the vehicle collides with pedestrian without braking, and the parameters should meet:

$$\frac{x_1 - x_2}{v_2} > \frac{y + l}{v} > \frac{x_1 - x_2}{v_1}. \quad (14)$$

The pedestrian velocity ( $v$ ) in this study is set 2 m/s, the upper ( $v_1$ ) and lower ( $v_2$ ) limits of the velocity are 60 km/h and 20 km/h, the vehicle width ( $l$ ) is 2 m, and the longitudinal distance between the pedestrian and the vehicle ( $y$ ) is 1 m. Combining formula (14), comprehensively considering the time required for the participant to take over and operate, it is finally determined that the lateral distance between the vehicle and the pedestrian ( $x_1 - x_2$ ) is 16 m.

This study also assists the construction of dynamic and complex traffic scenes through logical reasoning and the use of vehicle kinematic models. For example, when a vehicle stops in front of a zebra crossing, blocking the sight of the following vehicle, the following vehicle changes lanes and overtakes the preceding vehicle and collides with the pedestrian walking on the zebra crossing, as shown in Figure 3.

**3.1.4. Procedure.** Before the experiment, participants were asked to sign an informed consent form and fill out a questionnaire. Then, the test personnel needed to introduce



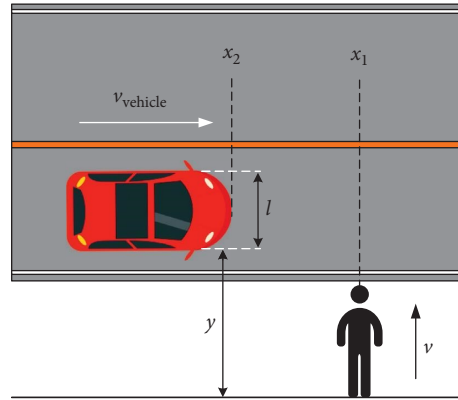


FIGURE 2: Dangerous traffic scene.

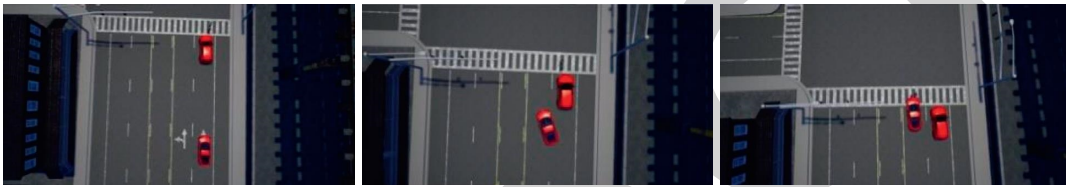


FIGURE 3: Complex dangerous scene.

the test content such as the switching operation of the automated driving system, the scenes of emergency take-over, and non-driving-related tasks to the participants. In addition, about 10 minutes of practice time should be left for the participants. The practice content includes being familiar with the sensitivity of the steering wheel, accelerator, and brake pedal, being acquainted with the experimental road environment, and practicing switching of automated driving modes.

During the experiment, the vehicle was driving along the calibrated line in automated driving status. The participants always performed NDRT, looking at the tablet computer, with both hands relaxed on both sides of the body, and their right foot relaxed and not placed on the brake or accelerator pedal. When encountering an emergency take-over situation, the automated driving system issued the take-over prompt sound of “please take over” mixed with buzzer and human voice according to the set TOR time. At this time, the participant pressed the switch button on the steering wheel to switch to manual driving mode and bypassed the broken down vehicle or obstacle in front. After that, the participants drove the vehicle back to the middle lane as soon as possible and switched to the automated driving mode.

**3.1.5. Data Collection.** The data used in this article are collected by the warning assistance system supporting the CARLA driving simulation platform. The collected parameters include the lane departure, vehicle velocity, acceleration, heading angle, pitch angle, and tilt angle. The left and right data of the steering wheel, brake data, front-wheel angle data, throttle depth, handbrake status, and gear status of the driving simulator are shown in Table 3.

The throttle depth and brake depth are both represented by 0 or 1. The various statuses of gears are 1, 2, 3, and reverse (R). The handbrake status has two states: “Yes (Y)” and “No (N).”

**3.2. Data Preprocessing.** In Figure 4(a), the black line represents the actual displacement trajectory drawn according to the vehicle position information and the blue line represents the target trajectory drawn by the position information obtained by the GPS sensor. The red line indicates the GPS target trajectory filtered by the Kalman filter. X-direction position refers to the distance that the vehicle swings laterally, and y-direction position means the vertical swing distance of the vehicle during automated driving. It can be seen from the figure that due to the interference of noise, some position measurement values obtained by GPS have a comparatively larger slice offset than the actual value, but the Kalman filter effectively filters the interference of noise, making the filtered target trajectory becomes smoother and more closely to fit the actual displacement trajectory. The position deviation of each measurement point and the actual point filtered by the Kalman filter is shown in Figure 4(b), and the deviations fluctuate within 0–10 meters.

The result of using the BRB method to verify the velocity and acceleration of the vehicle is shown in Figure 5. It can be seen that the overall confidence level is maintained at a relatively high level before the 11th second but showed a significant downward trend after the 11th second. This is because the motion model of the vehicle in the simulation adopts the constant acceleration model. At the beginning of the simulation, the vehicle speed gradually exceeds the set

TABLE 3: Data collection information sheet.

Sources	Name	Unit	Frequency (Hz)	Attribute
CARLA simulation platform	Velocity	km/h, m/s	60	Continuous
	Angular velocity	rad/s	60	Continuous
	Acceleration	$m/s^2$	60	Continuous
	Vehicle position*	M	60	Continuous
	Lane offset	M	60	Continuous
	Front-wheel angle	$^\circ$	60	Continuous
Driving simulator	Throttle depth	None	60	Continuous
	Brake depth	None	60	Continuous
	Handbrake status	None	-	-
	Steering wheel left and right corner	$^\circ$	60	Continuous
	Gear status	None	-	-

\*The lateral, vertical, and height positions of the vehicle are derived from the positioning system that comes with the CARLA simulation platform.

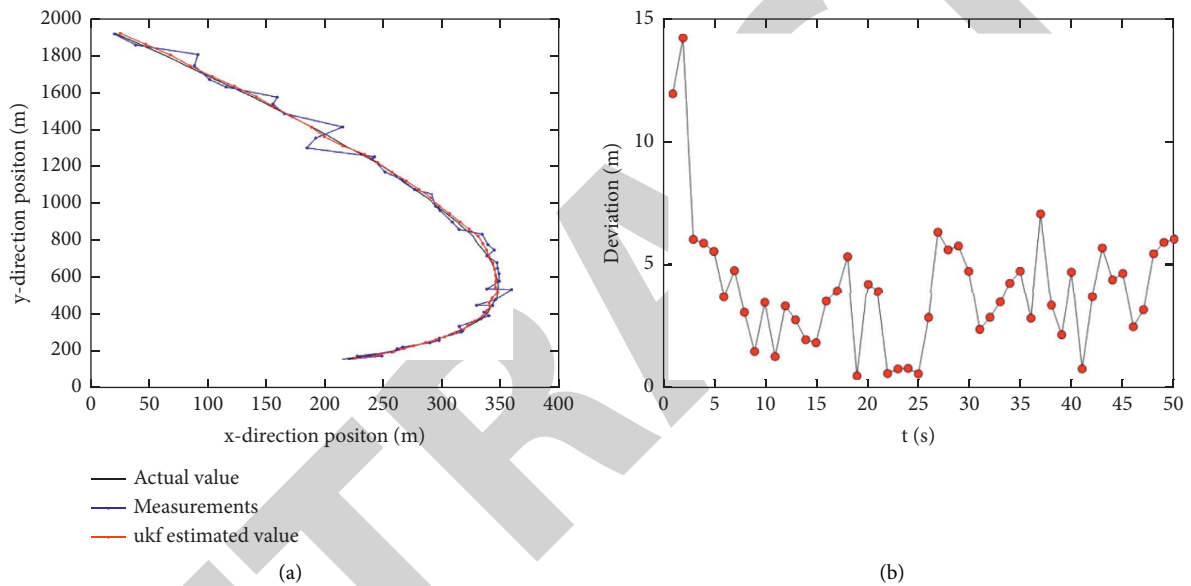


FIGURE 4: Kalman filter trajectory comparison (a) and the position deviation of each measurement point and the actual point after KF (b).

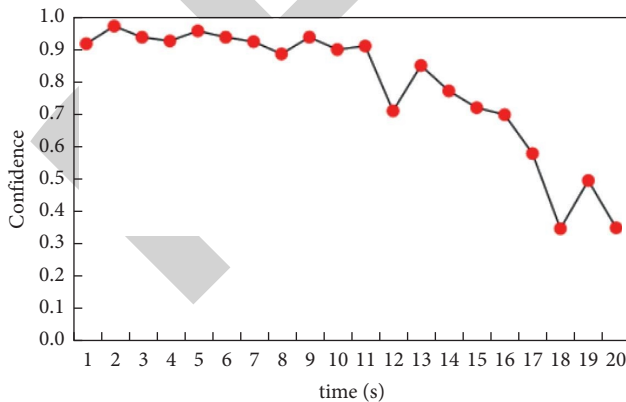


FIGURE 5: BRB verification confidence distribution.

range after driving for a short period of time. The BRB method judges that the simulated vehicle behavior is inconsistent with the conventional vehicle behavior and prompts that the vehicle is behaving abnormally at the 12th

second, so the confidence level continued to decrease. Overall, it is satisfactory that the target trajectory is effectively achieved by the verification of the simple kinematic model combined with the Kalman filter and the BRB method based on vehicle velocity and acceleration.

**3.3. Performance Evaluation.** In the low take-over frequency experiment, the vehicles driven by participants No.1, No.2, No.7, No.8, No.9, No.11, and No.14 on the driving simulator collided with obstacles, resulting the corresponding take-over failures. Therefore, in the descriptive statistics in Figures 6–9, the low-frequency take-over items do not have the data numbered above.

**3.3.1. Performance Evaluation Based on Experimental Parameters.** Figure 6 is a statistical histogram of the average time to complete take-over tasks of participants in the three experiments. It can be seen that as the frequency of take-overs increases, the average time for most participants to

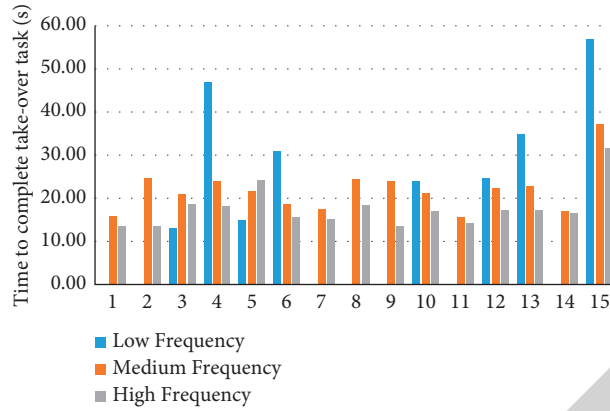


FIGURE 6: Average time to complete take-over tasks in each experiment of all participants.

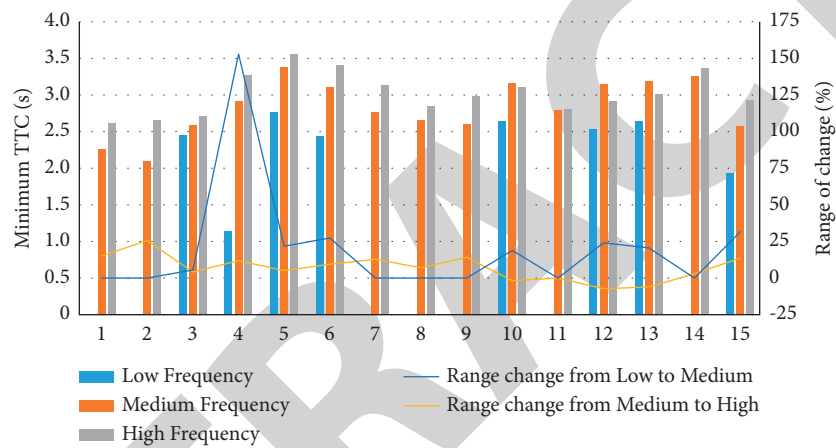


FIGURE 7: Average minimum TTC and change range in each experiment of all participants.

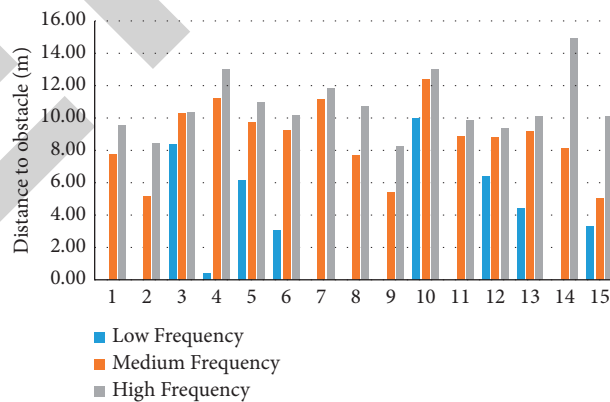


FIGURE 8: Average distance to obstacle in each experiment of all participants.

complete the tasks decreases. In the high-frequency take-over task, the participants performing the subtasks are more focused on completing the driving tasks, which are more efficient than the low-frequency and medium-frequency take-overs. However, there are a small number of participants who become cautious as the frequency of take-overs increases, resulting in a decrease in the efficiency of

completing tasks. For example, No. 3 and No. 5 participants are more relaxed under the low- and medium-frequency take-over intensity, and the complete efficiency of task is higher.

The minimum TTC reflects the risk acceptance level of participants during deceleration. The participants who crashed in the low-frequency experiments compare the

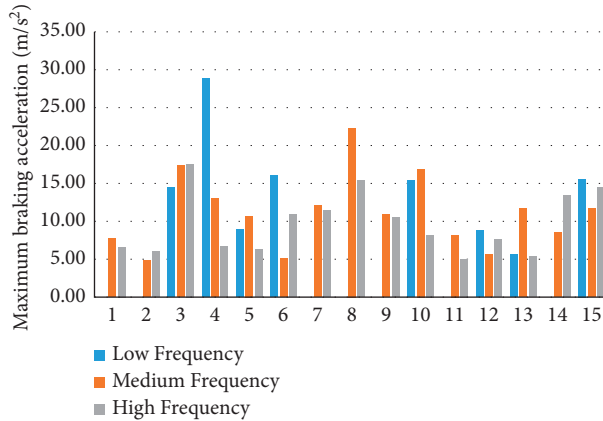


FIGURE 9: Average maximum braking acceleration in each experiment of all participants.

high-frequency data with the medium-frequency data, and other participants compare the high-frequency data with the low-frequency data. It can be seen from Figure 7 that the minimum TTC average value of all participants increases to varying degrees as the frequency of take-overs increases, which means that the risk awareness of participants has increased under the condition of increasing take-over frequency.

Due to the experiments excluding the minimum TTC data of the participants who collided in the low take-over frequency test, the variation range of some participants from low to medium take-over frequency is 0. It is not difficult to see that except for the participants who collided in the low take-over frequency, the minimum TTC change range of most participants from low to medium take-over frequency is significantly higher than the change range from medium to high take-over frequency. This shows that moderately increasing the take-over frequency can significantly improve the take-over effect of the participant, which is conducive to the concentration of participants. At the same time, the high take-over frequency does not significantly improve the participant's take-over performance compared with the medium take-over frequency.

The statistical results of the average distance to obstacle in the three experiments of all participants are shown in Figure 8. The driving style of each participant is different and there are differences between individuals, but with the increase in the take-over frequency, the average distance to obstacle of all participants increases to varying degrees. Among them, No. 8 and No. 14 increased significantly, which may be that the collision with obstacles in the low-frequency experiment makes them more focused on the take-over task in the high-frequency experiment. It shows that the increase in high frequency is more significant than that of low frequency. Therefore, it can be concluded that after the take-over frequency is increased from low to medium, the minimum TTC of the participant increases significantly, but if the take-over frequency continues to increase, the range of change is not obvious. In addition, there are high-frequency data of Nos. 2, 9, and 15 that have the same performance. Therefore, collisions cannot be used

as all the explanatory factors for the higher data in the participant's high-frequency experiment.

It can be seen from Figure 9 that the average maximum braking acceleration during the take-over process does not change regularly, which shows that the increase in the take-over frequency has little effect on the participant's braking behavior during the take-over process.

*3.3.2. Differentiated Performance Evaluation Based on Experimental Parameters.* Under the task of watching the video, the 5 s take-over request time can ensure the safety of take-over, but under the condition of low take-over frequency, 7 of 15 participants had a collision event, where 6 are males and 1 is female. The proportion of collision events among 9 males is 66.67%, while the proportion of 6 females is only 16.67%, and the total collision event rate is 46.67%. This shows that under the condition of low take-over frequency, even if the secondary tasks being performed by the participants and the take-over request time can ensure the safety of take-over, the probability of a collision event is still very high, and the participant is in a dangerous driving state. The collision probability of male participants is significantly higher than that of female participants under low take-over frequency, indicating that females have better risk awareness than males when performing take-over tasks. It can ensure the safety of the participant's take-over under the conditions of medium and high take-over frequency.

The average minimum TTC and the average distance to obstacles for male and female participants at three different frequencies (excluding collision data) are shown in Figure 10. It can be seen that the changing trends of the two parameters are very similar. In the low-frequency take-over experiment, the two parameter values of female participants are lower than the overall average level, but the males are higher than the average level. This indicates that the performance of males at low take-over frequency is severely polarized, and the take-over performance that can complete the take-over task is better, while the rest of the male participants have a collision. With the increase in take-over frequency, the two parameters of male and female have increased. Among them, the change in males is relatively gentle, which is approximately a low-slope linear shape. When the frequency of female take-over increases from low to medium, the increase in the parameters is larger and exceeds the average level. When the frequency of female take-over increases from medium to high, the increase is small, and the average level is slightly higher than that of males. Considering that there are more collisions among males at low frequency, females outperform males on average speed in the overall task.

The average speed of the male and female tasks is shown in Figure 11. The average speed in the task is the average speed of the participant during the period from taking over to the end of switching to automated driving again. It can be seen that under the conditions of low and medium frequency, the average driving speed of females is higher than that of males, and the average speed of both is significantly improved when the frequency changes from low to medium.

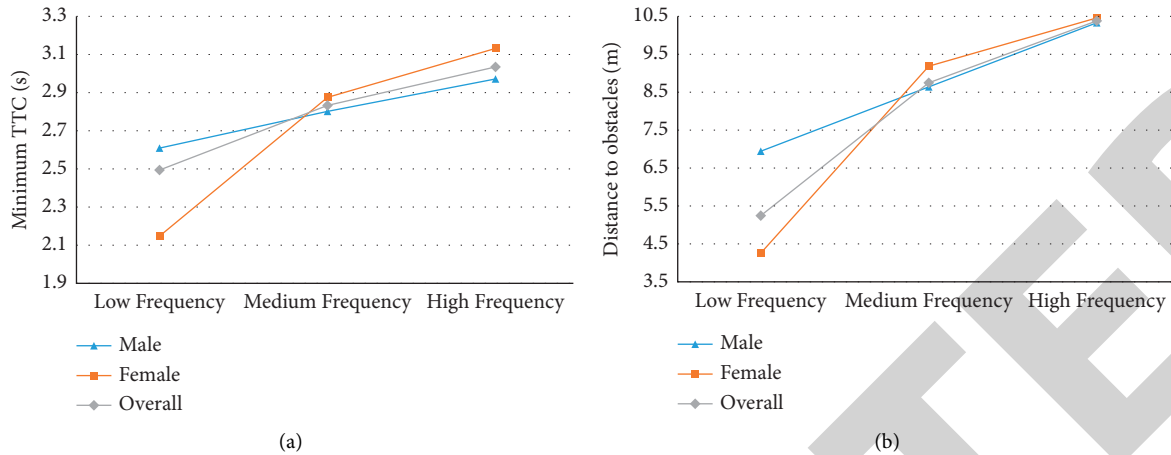


FIGURE 10: Average minimum TTC (a) and distance to obstacle (b) in each experiment (excluding collision data) of male and female participants.

When the frequency changes from medium to high, the average speed of males still increases significantly, while that for females decreases slightly. This means that when the take-over frequency increases within a certain range, the average speed of both males and females increases to a certain extent. When this range is exceeded, females generally become very cautious and lead to decrease in the average speed of the task, but males continue to increase, as does the overall trend.

The time to complete task and maximum braking acceleration of the participant at three different frequencies are shown in Figure 12. It can be seen from (a) that with the increasing frequency of take-over, the time for both males and females to complete the take-over task has been decreasing, and the change in frequency of take-overs from low to medium is more significant, but the time for females to complete take-over tasks is generally higher than that of males, indicating that females are very cautious when performing take-over tasks. In terms of maximum braking acceleration (b), males have the highest value at medium take-over frequency while females have the highest value at low frequency. The value of both males and females is the same at medium frequency. It can be seen that the take-over performance of females is inferior to that of males at low take-over frequency, and the other two take-over frequencies are similar between males and females.

### 3.4. Binary Logistic Take-Over Safety Evaluation Model.

The binary logistic regression method is used to construct a take-over safety evaluation model. The study in [22] pointed out that the minimum TTC lower than 1s can be used as an effective way to evaluate collisions. The minimum TTC is selected as the dependent variable, and the minimum TTC of 1s should be used as the boundary between the dangerous group and the safety group. However, the test driving scenes concluded in the references are mostly urban road scenes, and the vehicle speed is 50 km/h. Due to the limitation of the simulated

map, the vehicle velocity is 20–24 km/h, and the average value is 22 km/h. Under the same safety distance, the minimum TTC division limit should be taken as  $50 \times 1/22 \approx 2.3$  s. Therefore, this study chooses the minimum TTC less than or equal to 2.3 s as the dangerous group and more than 2.3 s as the normal group. Input variables select the time from the alarm to the braking reaction time, the maximum braking acceleration to take over, the gender of the participant, the distance to obstacle, and the duration of a single take-over. The method of selecting variables is “Backward Stepwise Regression: Maximum Partial Likelihood Estimate Likelihood Ratio Test (LR),” which means that all independent variables are first entered into the equation and then removed by the LR test. The probability cutoff value is set to 0.5. When the predicted probability value is greater than 0.5, the classification prediction value of the explained variable is considered to be 1 (dangerous group), and when it is less than 0.5, the classification prediction value is considered to be 0 (safe group).

The inspection results of model fitting are shown in Tables 4 and 5. Table 4 shows the Omnibus test of model coefficients, in which the “model” line outputs the likelihood ratio inspection results of whether all parameters in the logistic regression model are 0. A significant  $P < 0.05$  indicates the variables included in the fitted model. Among them, the OR value of at least one variable is statistically significant, and the overall model is meaningful.

Table 5 shows the test results of Hosmer and Lemeshow, indicating the goodness-of-fit results of the test model. When the  $P$  value is not less than the inspection level (i.e.,  $P > 0.05$ ), it is considered that the information in the current data has been fully extracted, and the model has a high degree of goodness of fit.

The likelihood values and significance changes when the model removes variables are shown in Table 6. If the significant change when removing a variable is less than 0.05, it

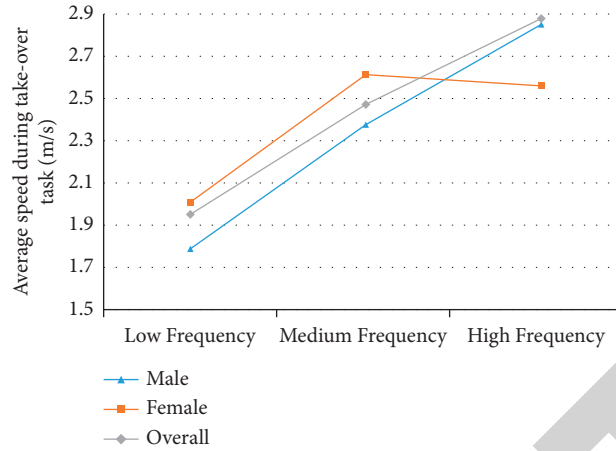


FIGURE 11: Average speed during the task in each experiment of male and female participants.

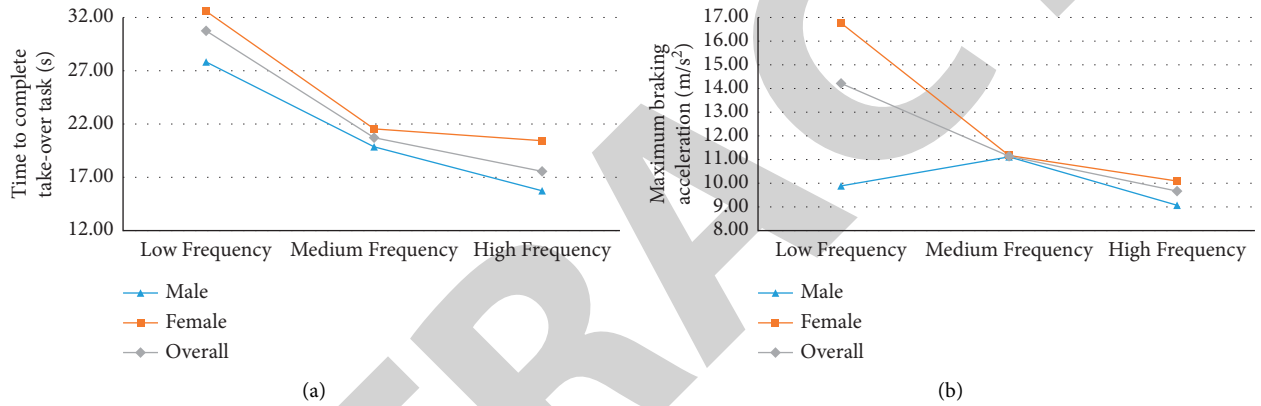


FIGURE 12: Time to complete take-over task (a) and maximum braking acceleration (b).

TABLE 4: Omnibus test of model coefficients.

		Chi-square	Degree of freedom	P value
Step 1	Step	23.974	5	0.000
	Block	23.974	5	0.000
	Model	23.974	5	0.000
Step 2	Step	-0.435	1	0.510
	Block	23.539	4	0.000
	Model	23.539	4	0.000
Step 3	Step	-1.731	1	0.188
	Block	21.808	3	0.000
	Model	21.808	3	0.000

TABLE 5: Hosmer and Lemeshow test.

Step	Chi-square	Degree of freedom	P value
1	8.248	8	0.410
2	4.155	8	0.843
3	5.643	8	0.687

means that the item is significantly related to the model and cannot be removed. In step 1 and step 2, the reaction time and braking acceleration with a significant change greater

than 0.05 are removed. In step 3, the significant changes when all independent variables are removed are less than 0.05, so they can no longer be removed.



TABLE 6: Model changes when variables are removed.

	Independent variable	Model log-likelihood	-2 change in log-likelihood	Significant change
Step 1	Gender	-43.234	4.904	0.027
	Reaction time	-41.000	0.435	0.510
	Braking acceleration	-41.709	1.853	0.173
	Distance to obstacle	-44.021	6.477	0.011
	Take-over time	-42.793	4.023	0.045
Step 2	Gender	-43.756	5.513	0.019
	Braking acceleration	-41.865	1.731	0.188
	Distance to obstacle	-46.519	11.038	0.001
	Take-over time	-43.279	4.558	0.033
Step 3	Gender	-44.652	5.573	0.018
	Distance to obstacle	-46.657	9.583	0.002
	Take-over time	-43.714	3.698	0.049

TABLE 7: Variables in the equation.

	Independent variable	Partial regression coefficient B	Standard deviation of error (S.E)	Wald test	P value	Exp (B)
Step 1	Gender	-1.592	0.817	3.796	0.051	0.203
	Reaction time	-0.645	0.974	0.439	0.508	0.525
	Braking acceleration	-0.084	0.060	1.939	0.164	0.919
	Distance to obstacle	0.282	0.117	5.865	0.015	1.326
	Take-over time	0.168	0.093	3.237	0.072	1.182
	Constant	-0.384	3.343	0.013	0.909	0.681
Step 2	Gender	-1.663	0.812	4.191	0.041	0.189
	Braking acceleration	-0.082	0.061	1.808	0.179	0.922
	Distance to obstacle	0.316	0.105	9.144	0.002	1.372
	Take-over time	0.177	0.093	3.665	0.056	1.194
	Constant	-2.166	1.994	1.180	0.277	0.115
Step 3	Gender	-1.663	.810	4.214	0.040	0.190
	Distance to obstacle	0.282	0.098	8.236	0.004	1.326
	Take-over time	0.154	0.088	3.090	0.079	1.166
	Constant	-2.243	1.925	1.357	0.244	0.106

The variable coefficients in the model regression process and their influence on the model are shown in Table 7. It can be seen that in the final model, gender and the distance to obstacle have a greater impact on the model.  $P(\text{gender}) = P(\text{distance to obstacle})$ , but  $\text{Wals}(\text{gender}) > \text{Wals}(\text{distance to obstacle})$ , and  $\text{Exp}(B)(\text{gender}) > \text{Exp}(B)(\text{distance to obstacle})$ , so the distance to obstacle is the variable that has the greatest impact on the model. After the simulation of the binary logistic model, it is concluded that the distance to obstacle is the variable that has the greatest impact on the participant's take-over performance. Among all the selected variables, the distance to obstacle is the variable that can most directly reflect the participant's take-over effect, which is consistent with the facts, and indicates that the results of model fitting are consistent with normal logical judgments.

In this model, females are set to 0, males are set to 1, and gender is the second most influential independent variable. Its regression coefficient B is negative, indicating that the female's take-over performance is poor. This is because the collision data are used in the model data. Excluding, in 3.3.2, the statistical results show that the number of female collisions is significantly less than that of males. However, in this model, taking the minimum TTC as the criterion for

take-over performance, it can only be concluded that the reaction time and operating time required by female participants are higher than those of male participants without a collision, which does not mean that male participants have significantly better take-over performance than female participants.

The final logistic regression equation is shown in equation (15), the model prediction accuracy rate is 87.7%, and the logistic classification table is shown in Table 8.

$$\text{Logit}(p) = -2.243 - 1.663x_1 + 0.282x_2 + 0.154x_3, \quad (15)$$

where  $x_1$  is the gender,  $x_2$  is the distance to obstacle, and  $x_3$  is the time taken for a single take-over.

Considering the effects of gender, reaction time, braking acceleration, distance to obstacle, and take-over time, the binary logistic model is applied to evaluate the take-over safety of L3 automated vehicles. The model is established with the minimum TTC as the dependent variable, and the Omnibus test and Hosmer and Lemeshow test are performed on the model coefficients to verify the validity and high fit of the model. The fitting results of the model are consistent with normal logic, and the prediction accuracy



TABLE 8: Logistic classification table.

	Minimum TTC $\leq 2.3$ s	Minimum TTC $> 2.3$ s	Percentage correction (%)
Minimum TTC $\leq 2.3$ s	5	14	26.3
Minimum TTC $> 2.3$ s	1	102	99
Total percentage (%)		87.7	

rate of the model is 87.7%, which is highly reliable. According to the analysis of the model, the variable that has the greatest influence on the participant's take-over performance is the distance to obstacle, which is also the variable that most directly reflects the participant's take-over effect among all the variables selected. In addition, gender, as the second most influential variable predicted by the model, also occupies a large proportion in performance analysis and safety evaluation. The reaction time, braking acceleration, and take-over time are also significantly related to take-over performance through model verification. Therefore, these factors are also indispensable when evaluating take-over performance and reflect the take-over effect to a certain extent. The influence mechanism of the above factors is complex and multidimensional.

According to the results obtained by the model, looking at the distance data of the participant to the obstacle in Figure 10(b), the take-over effect of males at low take-over frequency is significantly better than that of females, but the take-over frequency at medium and high take-over frequencies is roughly the same. Based on the evaluation of take-over performance in 3.3, both males and females perform better in the frequency of take-overs, which provides a reference for the study of the safety of conditionally automated driving.

#### 4. Discussion

This study studies the impact of take-over operations on take-over performance when the participant performs NDRT under L3 automated driving and evaluates the take-over safety of the experiment. To this end, by carrying out a simulation experiment on a driving simulator, collecting relevant driving data, and using the processed data to evaluate the take-over effect of the participant. Finally, the safety of take-over is evaluated based on the binary logistic model. With the deepening of the research, the results show that various factors affect the performance of the take-over, and the reasonable mode of control switching between the automated driving system and the driver needs to be further explored. This research promotes the research of data preprocessing methods, explores the research on take-over performance in L3 automated driving, and enriches the research on take-over safety.

A data preprocessing method based on the Kalman filter and BRB method is proposed. The target trajectory filtered by the Kalman filter is smoother and more suitable for the actual displacement trajectory. The BRB method associated with vehicle velocity and acceleration data is used as a trajectory tracking method to check the target position with high reliability and can effectively calibrate the abnormal behavior of the vehicle. The use of the Kalman filter is consistent with the current method of eliminating

interference in the literature [23], but more and more literature [24] combines the Kalman filter and other methods to track the trajectory or use other methods to achieve target tracking [25].

Descriptive and differential analysis of data such as the average time to complete the task, the minimum TTC, the distance to obstacle, and the maximum average braking acceleration shows that the take-over effect of females at low take-over frequency is weaker than that of males. However, the take-over performance of the medium and high frequencies is roughly equivalent. Due to males having too many collisions at low take-over frequency, the overall take-over performance of females is better than that of males. At present, most documents [26] are statistical analysis based on data, but there are also documents [27] that evaluate the take-over performance by constructing a structural equation model to provide new ideas for performance evaluation.

When evaluating the safety of take-over, a binary logistic model is applied to find out the variables that have a greater impact on the participant's take-over performance and further evaluate the safety of take-over based on these variables. Most of the take-over studies of L3 automated vehicles are based on test data for descriptive and differential analysis to evaluate take-over performance. In take-over-related research, take-over performance evaluation is mainly based on data [28]. At present, there are relatively few documents that determine the more significant variables through prediction.

#### 5. Conclusion

This study designs an urban road take-over scenario in a dangerous scenario based on a driving simulator, uses the data preprocessing method of the Kalman filter and BRB method, analyzes the impact of NDRT on take-over performance in L3 automated driving, and establishes a take-over safety evaluation model, and the specific conclusions are as follows.

The data preprocessing methods of the Kalman filter and BRB method have smaller deviations, are closer to the real trajectory, have certain reliability, and can effectively calibrate the abnormal behavior of the vehicle.

The overall driving performance of the participant improves as the frequency of take-overs increases. Under low take-over frequency conditions, it may be necessary to take appropriate measures to prevent the participant from focusing on the secondary task for a long time, such as intermittently reminding the participant to pay attention to the road conditions ahead, limiting the time for the participant to perform secondary tasks, etc. The overall driving performance of females is higher than that of males.

The binary logistic model uses the minimum TTC as the dependent variable to analyze that the variable that has the greatest impact on the take-over performance of participants is the distance to the obstacle, and gender is the second most influential variable. The prediction accuracy of the model is 87.7%, which has a high degree of credibility and validity.

## Data Availability

The datasets used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this study.

## Acknowledgments

This research is supported by the National Nature Science Foundation of China (52162049, 51805169, 52062014, and 52062015) and Natural Science Foundation of Jiangxi Province (20202BABL212009). This research is also jointly supported by Jiangxi Provincial Major Science and Technology Project-5G Research Project (Grant no. 20212ABC03A07).

## References

- [1] China's 11 departments jointly issued, "Smart Car Innovation and Development Strategy," *Robot Technology and Application*, vol. 2, p. 1, 2020.
- [2] M. L. Cunningham, M. A. Regan, T. Horberry, K. Weeratunga, and V. Dixit, "Public opinion about automated vehicles in Australia: results from a large-scale national survey," *Transportation Research Part A: Policy and Practice*, vol. 129, pp. 1–18, 2019.
- [3] K. F. Yuen, Y. D. Wong, F. Ma, and X. Wang, "The determinants of public acceptance of autonomous vehicles: an innovation diffusion perspective," *Journal of Cleaner Production*, vol. 270, Article ID 121904, 2020.
- [4] T. Zhang, D. Tao, X. Qu et al., "Automated vehicle acceptance in China: social influence and initial trust are key determinants," *Transportation Research Part C: Emerging Technologies*, vol. 112, pp. 220–233, 2020.
- [5] S. Haghzare, J. L. Campos, K. Bak, and A. Mihailidis, "Older adults' acceptance of fully automated vehicles: effects of exposure, driving style, age, and driving conditions," *Accident Analysis & Prevention*, vol. 150, Article ID 105919, 2021.
- [6] L. Wang, H. Zhong, W. Ma, M. Abdel-Aty, and J. Park, "How many crashes can connected vehicle and automated vehicle technologies prevent: a meta-analysis," *Accident Analysis & Prevention*, vol. 136, Article ID 105299, 2020.
- [7] G. Xiao, J. Lee, Q. Jiang, H. Huang, M. Abdel-Aty, and L. Wang, "Safety improvements by intelligent connected vehicle technologies: a meta-analysis considering market penetration rates," *Accident Analysis & Prevention*, vol. 159, Article ID 106234, 2021.
- [8] S. H. Yoon, Y. W. Kim, and Y. G. Ji, "The effects of takeover request modalities on highly automated car control transitions," *Accident Analysis & Prevention*, vol. 123, pp. 150–158, 2019.
- [9] D. Chu, R. Wang, Y. Deng, L. Lu, and C. Wu, "Vibrotactile Take-Over Requests in Highly Automated Driving," in *Proceedings of the 2020 4th CAA International Conference on Vehicular Control and Intelligence*, Hangzhou, China, Dec. 2020.
- [10] S. Li, P. Blythe, Y. Zhang et al., "Should older people be considered a homogeneous group when interacting with level 3 automated vehicles?" *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 78, pp. 446–465, 2021.
- [11] Y. Wu, K. Kihara, K. Hasegawa et al., "Age-related differences in effects of non-driving related tasks on takeover performance in automated driving," *Journal of Safety Research*, vol. 72, pp. 231–238, 2020.
- [12] M. Klinge, J. Andersson, A. Habibovic, N. Nilsson, and A. Rydstrom, "Drivers' ability to engage in a non-driving related task while in automated driving mode in real traffic," *IEEE Access*, vol. 8, pp. 221654–221668, 2020.
- [13] P. Rauffet, A. Botzer, C. Chauvin, F. Said, and C. Tordet, "The relationship between level of engagement in a non-driving task and driver response time when taking control of an automated vehicle," *Cognition, Technology & Work*, vol. 22, no. 4, pp. 721–731, 2020.
- [14] Y.-K. Ou, W.-X. Huang, and C.-W. Fang, "Effects of different takeover request interfaces on takeover behavior and performance during conditionally automated driving," *Accident Analysis & Prevention*, vol. 162, Article ID 106425, 2021.
- [15] Z. Zhang and Z. Zhu, *LONGITUDINAL STUDY OF DRIVER BEHAVIOR CHANGE IN AUTOMATED VEHICLES OVER TIME*, University, Missouri, 2020.
- [16] A. Calvi, F. D'Amico, L. B. Ciampoli, and C. Ferrante, "Evaluation of driving performance after a transition from automated to manual control: a driving simulator study," *Transportation Research Procedia*, vol. 45, pp. 755–762, 2020.
- [17] S. H. Yoon and Y. G. Ji, "Non-driving-related tasks, workload, and takeover performance in highly automated driving contexts," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 60, pp. 620–631, 2019.
- [18] F. Zhou, X. Yang, and J. D. Winter, "Using eye-tracking data to predict situation awareness in real time during takeover transitions in conditionally automated driving," *IEEE Transactions on Intelligent Transportation Systems*, vol. 99, pp. 1–12, 2021.
- [19] K. Wiedemann, F. Naujoks, J. Wörle, R. Kenntner-Mabiala, Y. Kaussner, and A. Neukum, "Effect of different alcohol levels on take-over performance in conditionally automated driving," *Accident Analysis & Prevention*, vol. 115, pp. 89–97, 2018.
- [20] N. Du, F. Zhou, E. M. Pulver et al., "Predicting driver takeover performance in conditionally automated driving," *Accident Analysis & Prevention*, vol. 148, Article ID 105748, 2020.
- [21] Q. Lin, Z. Wang, and G. Lu, "Takeover safety evaluation model for L3 autonomous vehicles," *Automotive Engineering*, vol. 41, pp. 1258–1264, 2019.
- [22] W. Young, A. Sobhani, M. G. Lenné, and M. Sarvi, "Simulation of safety: a review of the state of the art in road safety simulation modelling," *Accident Analysis & Prevention*, vol. 66, pp. 89–103, 2014.
- [23] H. Azevedo, S. Kumaar, and C. Esterwood, *Real-Time Estimation of Drivers' Trust in Automated Driving Systems*, Social Science Electronic Publishing, Rochester, NY, 2020.
- [24] W. Li, H. Li, K. Xu, Z. Huang, K. Li, and H. Du, "Estimation of vehicle dynamic parameters based on the two-stage estimation method," *Sensors*, vol. 21, no. 11, p. 3711, 2021.

## Retraction

# Retracted: An Annealing Model Analysis and Research on the In of Community Public Sports on Economic Development

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] Z. Guo, "An Annealing Model Analysis and Research on the In of Community Public Sports on Economic Development," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3129638, 9 pages, 2022.

## Research Article

# An Annealing Model Analysis and Research on the In of Community Public Sports on Economic Development

Zhen Guo 

*School of Physical Education, Shaoyang University, Shaoyang 422000, China*

Correspondence should be addressed to Zhen Guo; 2353@hnsyu.edu.cn

Received 30 January 2022; Revised 17 March 2022; Accepted 18 March 2022; Published 15 April 2022

Academic Editor: Daqing Gong

Copyright © 2022 Zhen Guo. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The state vigorously develops the application of community sports and takes it as an important unit of social and economic development. With the development of the economy, residents' living standards are getting better, which also greatly promotes the development of community public sports. Community public sports have made great progress in terms of scale, hardware, and services, and their impact on society is getting worse. Come bigger. However, there are still some deficiencies in the development of community sports. How to accurately evaluate and predict its facilities, construction, and service levels and put forward corresponding development strategies is an urgent problem to be solved in the field of community public sports. In this paper, an iterative simulated annealing algorithm is proposed to evaluate the level of public sports services and help community public sports institutions identify their own development deficiencies. Through the simulation analysis of the community sports organizations, community members, venue facilities and funds, venue facilities and funds, administrators, and the community sports activities, it is found that the identification accuracy of the evaluation results of the five indicators is above 95%, while the initial threshold of community public sports and social and economic development and the prediction accuracy required by the standard of development status are both 10%. At the same time, the recognition time of the evaluation results in the five indicators is less than 7 seconds. In community public sports development trend prediction and service level judgment, the identification time of the evaluation result is less than 4 seconds, and the overall judgment result is better.

## 1. Introduction

With the development of digitization and computers, public sports and other related industries have developed rapidly, showing the characteristics of rapidity, diversity, and intelligence shown by Sadegh Barkhordari and Tehranizadeh [1]. Community sports mainly refer to the regional mass sports carried out nearby in a certain area where people live together, taking the natural environment and sports facilities of the jurisdiction as the material basis, taking all community members as the main body, and taking meeting the sports needs of community members, increasing the physical and mental health of community members, and consolidating and developing community feelings as the main purpose. Community sports have six elements: community sports organizations, community members, venues, facilities and funds, administrators, instructors, and community sports

activities. Public sports are affected by external and internal factors and cannot be accurately evaluated and analyzed by traditional data analysis methods, which affects the evaluation effect of service level (Bourliva et al. [2]) and limits the development of intelligent public sports. Some scholars believe that the purpose of the intelligent development of community public sports is to meet the development needs of the industry. Therefore, starting from the prediction of development trend, operation status, and the development of community public sports (Donoso [3]), a comprehensive evaluation of the public sports service level can provide a reference for the formulation of a service strategy. Some scholars also believe that the accuracy of intelligent evaluation in community public sports is low, and the proportion of unstructured data is high (Durmus [4]).

We can improve the accuracy of intelligent evaluation of public sports and reduce the interference of unstructured

data by extracting eigenvalues. Based on the above background, this paper uses metropolis acceptance criteria and a simulated annealing algorithm to analyze the eigenvalues and analyzes the service level of community public sports from the perspective of intelligence so as to improve the evaluation accuracy of public sports services. Based on the above reasons, this paper mainly introduces it from three aspects. Firstly, it describes the service level in the community of public sports mathematically, and introduces the intelligence of service level, the parameters of service level, and the processing of service level-related data. Secondly, a prediction model of public sports service level based on a simulated annealing algorithm is constructed. The constraints of the simulated annealing algorithm, the accuracy of public sports service level, and the prediction steps of public sports service level are analyzed. Thirdly, the simulation of public sports service level prediction model is based on a simulated annealing algorithm. We describe the cases of public sports services and analyze the accuracy and calculation time of prediction results.

From the above analysis, it can be seen that the rapid development of community public sports can promote the development of the social economy and make its decision-making develop in a positive direction. In order to better carry out the research of community public sports and make them play a role in economic promotion, it is necessary to introduce comprehensive judgment methods to improve the accuracy of calculation, which is of great practical significance. In addition, the comprehensive analysis of community public sports can also enhance the coupling between community public sports and economic development and provide support for development decision-making.

## 2. The Mathematical Description of Service Level Prediction in Community, Public Sports, and Social Economy under Simulated Annealing Algorithm

The iterative simulated annealing algorithm is a comprehensive calculation method, which can carry out iterative analysis on community public sports. This algorithm has the advantages of a simple calculation process and fast analysis speed. Compared with the classical simulated annealing algorithm, the iterative simulated annealing algorithm can carry out progressive calculation, eliminate redundant data, make the calculation result more accurate, and avoid local extremum.

*2.1. The Intelligent Description of Service Level.* Intelligent service level has two characteristics: on the one hand, it processes massive information, uncertain data, and diverse data, and the proportion of controllable data is >50%; on the other hand, multisource data are analyzed and dynamic information is obtained. The public sports data structure is as follows: the diversity of data sources, the comprehensiveness of public sports, and the diversity of calculation, observation, and measurement (Fitriingsih [5]), as shown in Figure 1.

The structure of public sports data is mainly in the form of controllable and uncontrollable factors and is combined with the economic data sources of public sports-related industries to form a large amount of public sports data. Public sports should start from the industry, conduct intelligent analysis in combination with industry development capacity, industry service status and trend, and improve the accuracy of public sports assessment.

It can be seen from Figure 2 that although the demand for sports for community public sports fluctuates, it is increasing on the whole. Therefore, community public sports play a vital role in the development of sports. Some scholars also believe that community public sports can also promote the development of a social economy. The specific contents are shown in Figure 3.

It can be seen from Figure 3 that there is a certain correlation between social public sports and the social economy, and the correlation is increasing year by year. In summary, community-based public sports play a supportive role in sport development and are somewhat correlated with social and economic development. In order to better carry out the development of community public sports and give play to their social promoting role, it is necessary to conduct in-depth research on community public sports. At the same time, depending on the current situation of public community sport, foresee its future development, in order to improve the development of sport and the social economy. As a result, a thorough analysis of community public sports is of great practical and theoretical importance.

*2.2. Description of Service Level Metrics.* The digitization of the level of public sports services and the diversification of the factors influencing demand increasingly highlight the characteristics of public sports data. The traditional level of service method cannot fully address the intelligent development needs of community-based public sports. Under the background of intelligent development, public sports prediction algorithms such as particle swarm optimization and support vector can meet the needs of large volume and multiple types and realize the prediction and analysis of public sports service level, community public sports service status, and public sports development trend. The parameters of the community public sports are set as follows (Haghverdi [6]): community sports organizations, community members, venue facilities and funds, administrators, directors, and coordinators of community sports activities.

- (1) Assume the community sports organizations are  $x_o$ , the community members are  $x_j$ , community outdoor fitness venues and funding are  $x_k$ , venue facilities and funds are  $x_b$ , administrators are  $x_m$ , the community sports activities is  $x_n$ , and the intelligent dataset of the community public sports service level is  $C$ . The result of the calculation shall be drawn up in accordance with the following formula:

$$C = \{c_1, c_2, \dots, c_i\}, \quad (1)$$

where  $c_i$  is shown in the following formula:

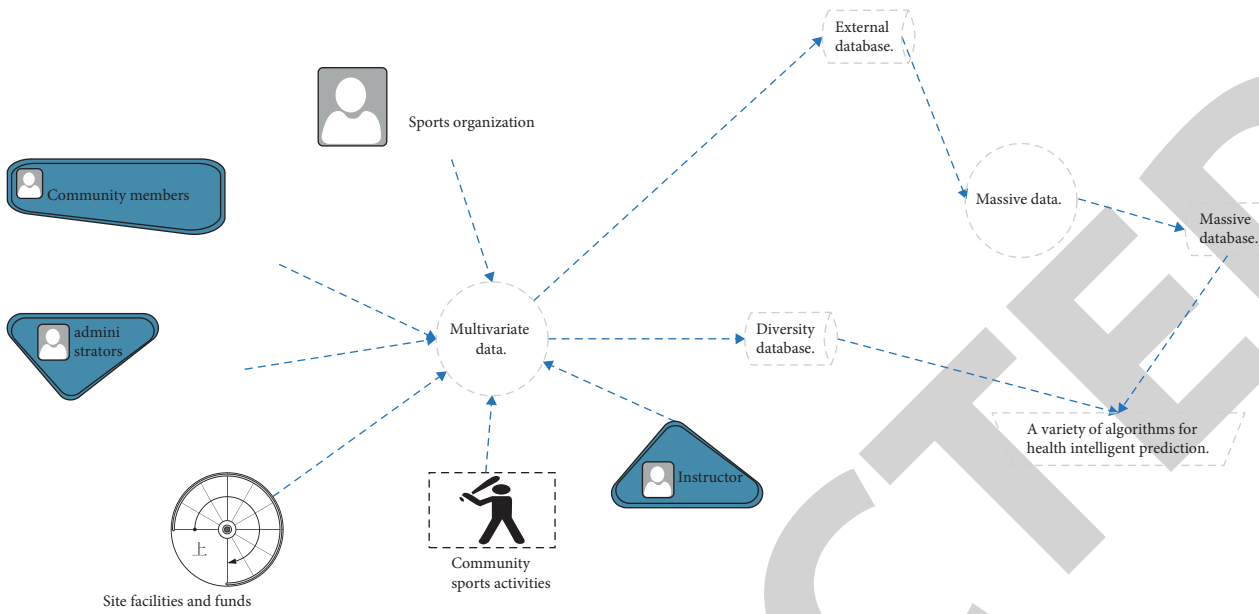


FIGURE 1: The data source of intelligent prediction of public sports in sports.

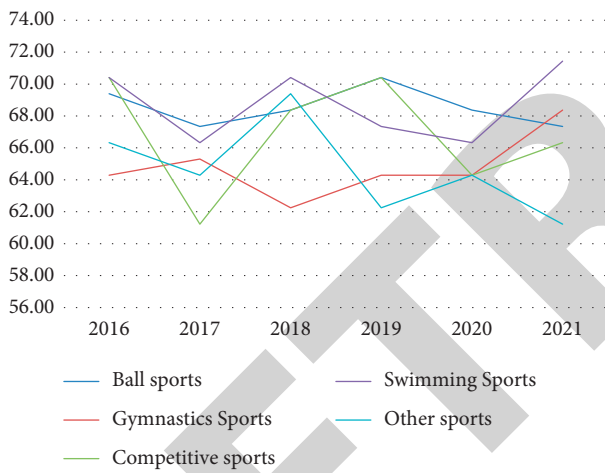


FIGURE 2: The demand for different sports for social public sports.

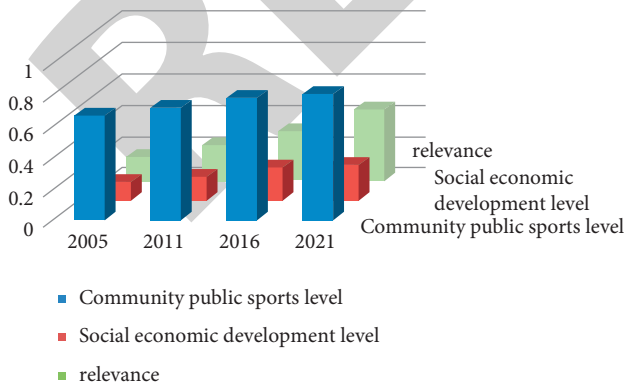


FIGURE 3: The correlation between community public sports and socioeconomic development.

$$c_i = \sum_{i,j,k}^n x_i + x_j + x_k + x_l + x_m + x_n, \quad (2)$$

where  $I, j, k, l, m, n$  belongs to natural numbers. The different data for the public service of sport come from different sectors.

- (2) Suppose  $n$  is an intelligent collection,  $i$  is the industry of data source,  $j$  is the structure type of data (uncontrollable factor data = 1, otherwise = 0),  $k$  is the data acquisition method (qualitative method acquisition = 1, quantitative method acquisition = 2, comprehensive acquisition method = 0),  $l$  is the intelligent development level (Note: 1~5 levels, the higher the value, the higher the level), and  $M$  is each data collection organization, then the data acquisition item can be described as  $N_{i,j,k,l,M,i,j,k,l,M} = (1, 2, \dots, n)$ , and  $n$  is a natural number.
- (3) Different public sports institutions divide similar industry sources, structure types, acquisition methods, and economic development level data into community public sports data sets.
- (4) Data from all data institutions are analyzed centrally by the Centre for Economic Analysis of Community Public Sport. Under the constraints of forecast accuracy, prediction time, and safety, the data obtained are classified according to the relevant weights and thresholds.
- (5) The collection time of health data and the intelligence level of health data in any given data organization are the same or similar (Han [7]).



- (6) For the stability of the data organization and analysis, the Fourier series method shall be adopted to reduce the interference of data organization, personnel, and

other nonresistance factors as shown in the following formula:

$$P(t) = a_0 + \sum_{i=1}^n \frac{[(y_i z_i u_i - y_{i-1} z_{i-1} u_{i-1})]}{2} + \frac{(y_i z_i u_i + y_{i-1} z_{i-1} u_{i-1})}{2} + \xi, \quad (3)$$

where  $P$  is the antiinterference ability described by the intelligent analysis of health data,  $t$  is the measured time of health data,  $y$  is the degree of intelligence,  $z$  is the interval of health data collection,  $u$  are the sports health standard, and  $\xi$  interference coefficient.

**2.3. Simplified Processing of Health Information Data.** The data gathered by the various public sports institutions must first be processed by K-means. The result is indicated in the form equation as follows:

$$\sum_{i=0}^n I_i = |S|. \quad (4)$$

Among them,  $S$  is the clustering range, the lower the value, the better the clustering effect.  $I$  is the data collected by the data organization, and  $I$  is the number of the data. In addition,  $S$  can be replaced by similarity  $J$  so as to improve the efficiency of the calculation. In order to simplify the process, the weight can be judged by using the frequency of a certain date, the data threshold collected by the data institution can be judged according to the previous historical data or the data frequency in the community public sports database, that is shown in the following formula:

$$w_i = \frac{H \cdot G_{ij}}{G_i + G_j}, \quad (5)$$

where  $G_{ij}$  is the number of times  $J$  data organization collects  $I$  data,  $G_i$  is the number of times  $I$  data appears in the community public sports database,  $G_j$  is the frequency  $j$  data organization sends data, and  $WI_i$  is the threshold of  $my$  data.

### 3. Construct a Public Sport Intelligence Prediction Model Based on a Simulated Annealing Algorithm

**3.1. The Limitations of the Simulated Annealing Algorithm.** Prior to the analysis of intersectoral data institutions in the community, public sports, all data institutions should be classified and judged according to applicable standards. In this paper, the metropolis acceptance criteria (Harkut et al. [8]) are adopted. If community-based public sports meet real needs, relevant data will be collected, otherwise, data organizing data will not be accepted. Metropolis accepts the judgment of the criteria and can save the resources of the community public sports, reduce the amount of data

processing, and improve the speed of intelligent development (Luo et al. [9]). Assuming that the state of the community public sports is  $the$ , we conduct metropolis acceptance criteria analysis to determine whether the state needs to be changed, that is, from  $my$  state to  $i+1$  intelligence. The calculation formula shall be

$$R(T) \begin{cases} R(T) = 0 \sim 0.5, \\ R(T) = 0.5 \sim 1. \end{cases} \quad (6)$$

The limitations of formula (2) appear in formula (7) as follows:

$$\begin{cases} f(i) \longrightarrow f(i), \\ f(i+1) \longrightarrow f(i+1), \\ \exp\left[\frac{f(i) - f(i+1)}{T_i}\right], \end{cases} \quad (7)$$

where  $\exp()$  is the expected function of service intelligence, which is to judge whether to be smart;  $R(T)$  is the result function of intelligence in community public sports. If the result of  $R(T)$  is  $0 \sim 0.5$ , it indicates that the degree of intelligence is low and the original state can be maintained. If the result ranges from  $0.5$  to  $1$ .

**3.2. The Appropriateness between Collective Sports and the Recut Algorithm.** The adjustment of communal public sports and the annealing algorithm is divided into (Luo [10]) local adjustment CP and global adjustment MP, which reflect the adjustment of the two under different aspects, which is also the premise of the calculation of the evaluation results. The larger the values of  $PC$  and  $Pm$ , the better the data fit. Since metropolis in 2.1 accepts the limitation of the criterion and requires both data to be optimized in order to become the calculation data, the calculation formula is as follows:

$$P_c \begin{cases} N_{1,1,0,1}, & R_c = R_{mg}, \\ \frac{\varphi_c(R_c - R_{\min})}{(R_{mg} - R_{\min})}, & R_c < R_{mg}, \end{cases} \quad (8)$$

$$P_m \begin{cases} N_{1,1,1,0}, & R_m = R_{mg}, \\ \frac{\varphi_m(R_m - R_{\min})}{(R_m - R_{\min})}, & R_m < R_{mg}, \end{cases} \quad (9)$$



where  $N_{1,1,0,1}$  is the initial value of  $P_c$ ,  $N_{1,1,1,0}$  is the initial value of  $P_m$ ,  $R_{\min}$  is the minimum value of the whole data,  $R_c$  is the local minimum value, and  $R_m$  is the overall minimum value.  $\varphi_c$  is the local fitting coefficient and  $\varphi_m$  is the overall fitting coefficient.

3.3. *The Accuracy of the Results for Predicting the Level of Public Sport Intelligence.* The corresponding datasets should be constructed before the prediction results are calculated. In

$$f(x \rightarrow y) = \min \left[ \int_{i,j}^n \sum_{i=1}^n \max(F(x_i), F(x_{i+1})), \sum_{j=1}^n \max(y_j), (y_{j+1}) \right] + \phi_i \sin(\xi_i), \quad (10)$$

where  $f(x \rightarrow y)$  is the relationship function between simulated annealing algorithm and community public sports, that is, the appropriate function, which reflects the matching degree between them. The maximum value of any two service data is as follows:

$$\sum_{i=1}^n \max(F(x_i), F(x_{i+1})). \quad (11)$$

The maximum value expected from any level of intelligence is as follows:

$$\sum_{j=1}^n \max(y_j), (y_{j+1}). \quad (12)$$

The level of information provided by public sports institutions to obtain the maximum or minimum of two points is as follows:

$$\int_{i,j}^n \left( \sum_{i=1}^n \max \left( F(x_i), F(x_{i+1}), \sum_{j=1}^n \max(y_j), (y_{j+1}) \right) \right). \quad (13)$$

The minimum of all extreme values is set out as follows:

$$\min \left[ \int_{i,j}^n \sum_{i=1}^n \max(F(x_i), F(x_{i+1})), \sum_{j=1}^n \max(y_j), (y_{j+1}) \right]. \quad (14)$$

3.4. *Steps in Predicting Community Public Sports Based on a Simulated Annealing Algorithm.* Using the above mathematical description, the following calculations are needed:

- (1) We set initial  $C = \{c_1, c_2, \dots, c_i\}$  values of community public sports, thresholds of data fitting (local fitting  $P_c$ , overall fitting  $P_m$ ) and calculation accuracy, intelligent prediction results  $N_i, j, k, l, m$  values, and initial metropolis acceptance criteria;
- (2) Gradient 200 iterative tests were performed on the fitted data (Reinaldi [13]);

this paper, the appropriate function is used to clarify the relationship between the simulated annealing algorithm and community public sports (Martin-Fernandez [11]). Under the conditions of preset accuracy and threshold constraints, the moderate function performs multiple iterative analyses, guides the operator to calculate in the positive direction, reduces the occurrence rate of local extreme values and "false eigenvalues" (Petroff [12]), and improves the accuracy of predictive results, as shown by the following formula:

- (3) We verify the calculation results according to metropolis acceptance criteria and constraints, incorporate the qualified results into the total scheme, and calculate the recognition accuracy and calculation time of the prediction results of the overall scheme (Ren e al. [14]);
- (4) We determine whether all the data we are going through will stop the calculation, otherwise proceed to step 3;
- (5) Finally, the minimum value should be taken out. The overall diagram and output of the computation process are shown. The specific process is presented in Figure 4.

## 4. Case Analysis of the Public Sports Information Prediction Model Based on a Simulated Annealing Algorithm

4.1. *The Case Introduction.* We take 4 provincial public sports establishments, 5 municipal public sports establishments, and 6 county public sports establishments from January 1, 2020, to December 2021 as examples. Song et al. [15] analyze the impact of the simulated annealing algorithm on the current state and trend of development of public sports services and determine the accuracy of the discrimination and time of the prediction results. The accuracy and calculation accuracy of data fitting (partial fitting PC and overall fitting PM) are set to 0.1 and the number of iterations is 200. Among them, the data collection is the statistical yearbook of public sports institutions, the website of public sports institutions, and a questionnaire survey (reliability and validity >0.7). The intelligence level is level 3 and follows the guidelines for the development status of community public sports (GB/T 31464) in 2015, The IP addresses of public sports institutions are 192.168.1.102~192.168.1.202.

4.2. *The Fit between Community Public Sports and Simulated Annealing Algorithm.* Through K-mean clustering and metropolis acceptance criteria, the data of intelligent and public sports industries are standardized, and 20 iterative

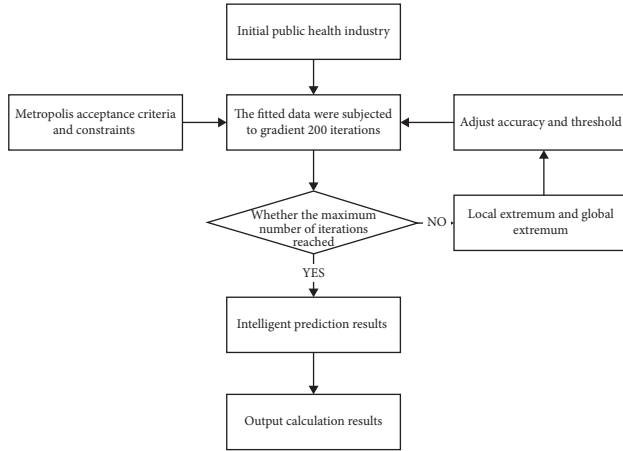


FIGURE 4: Community public sports analysis steps.

analyses are carried out to obtain the data fitting results, as shown in Table 1.

It can be seen from Table 1 that both local fitting and overall fitting are >95%, while the overall fitting is high and meets the threshold of 10% set by K-mean clustering and metropolis acceptance criteria. The results of the specific couplings are presented in Figure 5.

It can be seen from Figure 6 that there is a strong correlation between community public sports and social economy, which shows that community public sports present a large-scale and high-level development trend and can promote economic development and make it develop in a positive direction. At the same time, the rapid development of the economy also promotes the upgrading and development of community sports. For the classical simulated annealing algorithm, the iterative simulated annealing algorithm has a higher correlation, indicating that the correlation between social public sports and economic development is stronger, which lays the foundation for accurate calculation in the later stages.

#### 4.3. Community Public Sports and the Social Economy.

The results show that although the correlation between social public sport and social economy fluctuates in a small range, the overall correlation is high, between -20% and 180%. As a result, there is a strong relationship between community public sports and the social economy. The small fluctuation between community public sports and social economy is mainly determined by the whole society's sports policy and social-economic policy, and has nothing to do with the simulation results of the annealing algorithm in this paper. The specific outcomes are presented in Figure 6.

**4.4. The Accuracy of Prediction Results.** The community sports organizations are  $x_b$ , the community members are  $x_j$ , the venue facilities and funds are  $x_k$ , the venue facilities and funds are  $x_l$ , the administrators are  $x_m$ , and the community sports activities are  $x_n$ . The accuracy of the identification of predictive results in the five indicators directly determines the effectiveness of the entire model.

TABLE 1: Degree of adjustment of the various figures.

The adequacy of predictive data on the level of public athletic information	Iteration (times)	Fitting degree (%)
Local fitting	0~50	97.3
	50~100	98.7
Global fitting	0~50	97.2
	50~100	99.6
Total fitting degree	0~100	98.3

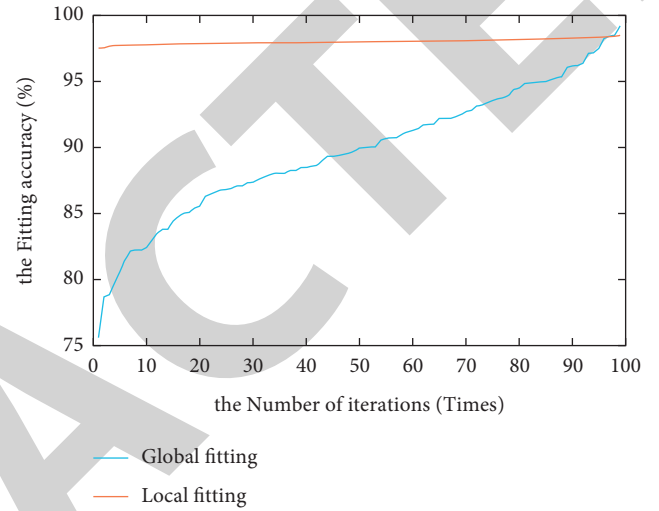


FIGURE 5: Results of local and global adjustment.

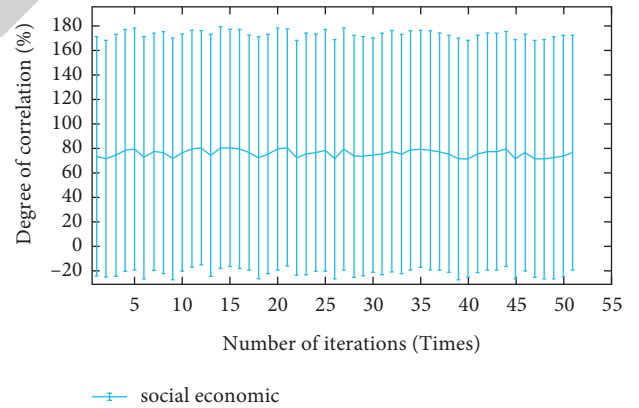


FIGURE 6: The correlation between community public sport and social economy.

Through MATLAB simulation analysis, it can be seen that the recognition accuracy of the prediction results in the abovementioned five indicators is >95%, while the initial threshold and the requirements in the criteria for the current service situation of community public sports are 0.1. At the same time, in the process of 40~50 iterations, the calculation accuracy changes greatly, which is mainly caused by the jumping change in social and economic development and does not affect the final calculation accuracy. So, the simulation results are better, and the results are shown in Table 2.

TABLE 2: Accuracy of forecast outcomes.

Judgment index	Iteration (times)	Fitting degree (%)
The community sports organizations	0~50	98.23
	50~100	97.23
The community members	0~50	96.62
	50~100	98.82
Venue facilities and funds	0~50	97.02
	50~100	98.32
Administrators	0~50	97.12
	50~100	98.24
The community sports activities	0~50	97.22
	50~100	97.12
Total fitting degree	0~100	98.62

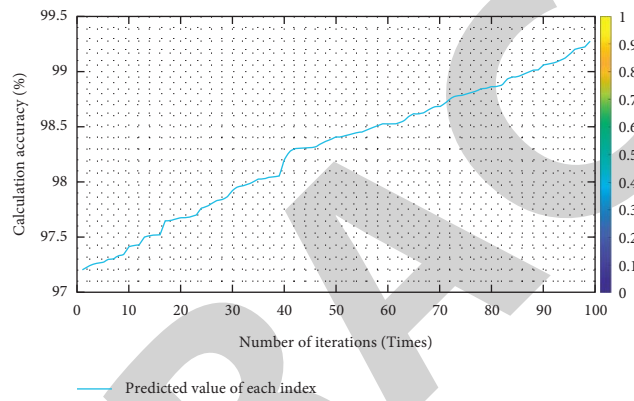


FIGURE 7: Accuracy of predictive results for each index.

The precision of the specific calculation can be derived from the data in Table 2, and the results are illustrated in Figures 7–9.

It can be seen from Figure 8 that the service judgment result of social public sports is better, which is between 75 and 99% as a whole, better than 82~89% of the classical simulated annealing algorithm. We can see that the computation result of the iterative simulation annealing algorithm is better.

It can be seen from Figure 9 that the economic development trend is relatively stable, but there are significant changes in 10~20 iterations and 50~60 iterations, mainly due to the changes in economic policies at that time. However, the accuracy of the overall evolution meets the demands.

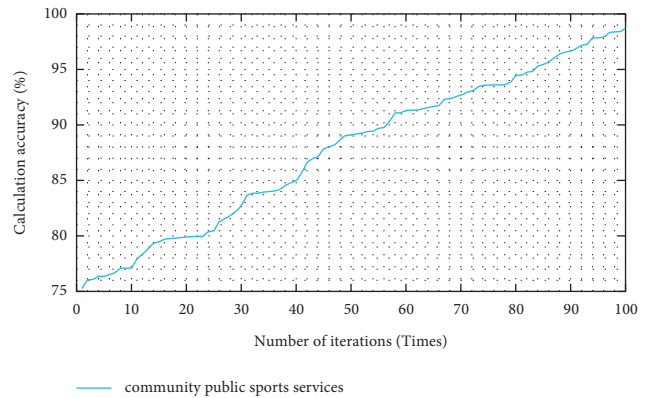


FIGURE 8: Judgment on the service status of communal public sports.

**4.5. Judgment Time of Prediction Results.** The characteristic of intelligence is that the data has a large amount of data processing and high processing complexity, so the identification time of prediction results is another verification index. The results show that the identification time of prediction results in community public sports service capability  $x_i$ , community public sports service status judgment  $x_j$ , and public sports service trend judgment  $x_k$  is less than 40 seconds. Compared with 0~10 iterations, the calculation time of 10~20 iterations is longer, mainly due to the large

amount of postprocessing data and the relatively complex calculation process. However, the identification time of the overall prediction results meets the requirements of the guidelines for the status of services in community public sports, and the results are shown in Table 3.

The overall calculation time of the level of public sports information can be derived from Table 3, as shown in Figure 10.

We can see in Figure 10 that the overall mounting time of the calculation result is better, between 0 and 1 s. That is, the

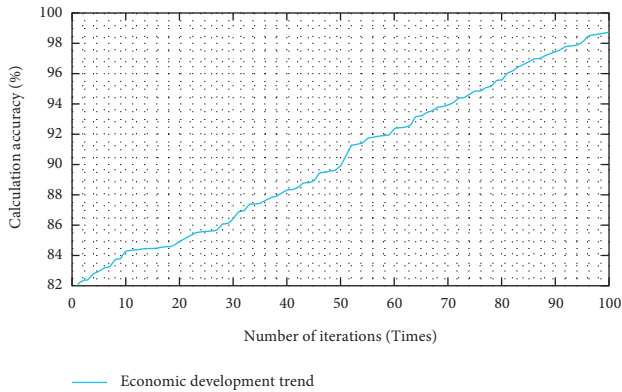


FIGURE 9: Outcome of the judgment trend in public sports services.

TABLE 3: Judgment time of forecast results.

Judgment index	Iteration (times)	Fitting degree (%)
The community sports organizations	0~10	$0.61 \pm 1.23$
	10~20	$0.53 \pm 0.92$
The community members	0~10	$0.32 \pm 2.01$
	10~20	$0.72 \pm 2.32$
Venue facilities and funds	0~10	$0.62 \pm 1.72$
	10~20	$0.72 \pm 0.93$
Administrators	0~10	$0.61 \pm 1.91$
	10~20	$0.73 \pm 0.12$
The community sports activities	0~10	$0.64 \pm 1.84$
	10~20	$0.71 \pm 0.82$
Total fitting time	0~20	$0.72 \pm 2.01$

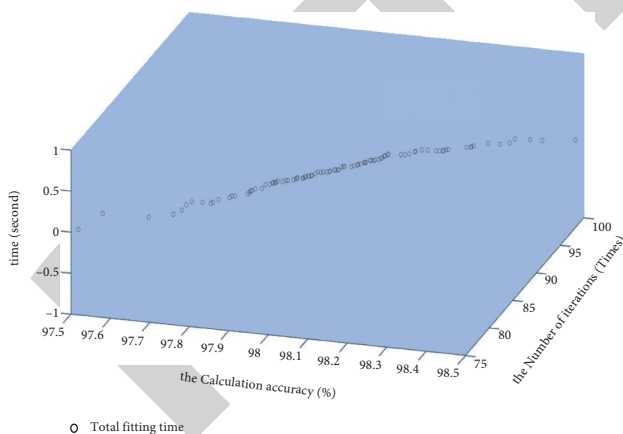


FIGURE 10: The total fitting time.

overall mounting time is short. At the same time, under different iteration and accuracy requirements, the fitting time has not changed, indicating that the overall calculation time is relatively stable, which further proves the effectiveness of this method. Compared to the traditional simulated annealing algorithm, the computational time of the iterative simulated annealing algorithm is shorter.

## 5. Conclusion

The social economy is the basis for the development of community public sports. With the improvement of the economic level, the development of community public sports is getting better and better. The development and improvement of community public sports promotes the development of the social economy to a large extent and have an inseparable relationship. The rapid development of the intelligence level of the community public sports makes the intelligence level higher and higher (Tubishat [16]), and the data volume and complexity show characteristics (Twahirwa and Wang [17, 18]), which makes traditional analytical methods unable to meet the analytical requirements of relevant data institutions. The studies of Yagmur and Kesen [19] have resulted in a significant decline in the accuracy of forecast results. A method of analysis based on a simulation annealing algorithm is proposed in this document [20]. Through the standardized processing of public sports service data, fitting data with an accuracy of 0.1 is obtained (Svalina et al. [20]). Then, using iterative calculation, the prediction results and identification rates of public sports service capability judgment  $x_i$ , public sports service status judgment  $x_j$ , public sports service trend judgment  $x_k$ , and other indicators are obtained by Zolfi and Jouzdani [21, 22]. The MATLAB simulation results show that the recognition accuracy of the prediction results in the five indicators is  $>95\%$  (Ilin [23]), while the initial threshold and the requirements in the guidelines for the development status of community public sports are 0.1. At the same time, the identification time of the prediction results in the five indicators is less than 7 seconds, and the identification time of the prediction results in the judgment of public sports service trend and intelligence level is less than 4 seconds (He and Ye [24]). So, the overall judgment result is better (Chen et al. [25]). The global adjustment time of the calculation result is better, between 0 and 1 s, i.e., the global adjustment time is short. At the same time, under different iteration and accuracy requirements, the fitting time has not changed, indicating that the overall calculation time is relatively stable, which further proves the effectiveness of this method. Compared with the classical simulated annealing algorithm, the calculation time of the iterative simulated annealing algorithm is shorter. However, there is still insufficient research on the coupling between public sports intelligence level indicators and intelligent judgment in this paper. In future research, we will focus on the analysis of the above contents.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported by Shaoyang University.

## Retraction

# Retracted: Research and Design of Docker Technology Based Authority Management System

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] L. You and H. Sun, "Research and Design of Docker Technology Based Authority Management System," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 5325694, 8 pages, 2022.



## Research Article

# Research and Design of Docker Technology Based Authority Management System

Li You<sup>1</sup> and Hui Sun<sup>2</sup> 

<sup>1</sup>Information Center, Hebei Vocational University of Industry and Technology, Shijiazhuang, China

<sup>2</sup>Department of Automotive Engineering, Hebei Vocational University of Industry and Technology, Shijiazhuang, China

Correspondence should be addressed to Hui Sun; [sunhui6033291@126.com](mailto:sunhui6033291@126.com)

Received 23 January 2022; Revised 24 February 2022; Accepted 25 February 2022; Published 13 April 2022

Academic Editor: Daqing Gong

Copyright © 2022 Li You and Hui Sun. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the development of mobile Internet technology and the continuous popularization of the network, various kinds of network software come out constantly and people are becoming more and more dependent on them, while role authority management is of great importance for the security of software, the control of management process, and the usability of users. In terms of system implementation, virtual machine technology is often faced with problems such as high virtualization overhead, poor scalability, and long deployment time in spite of its good isolation effect. Container technology represented by Docker can well solve these problems and make it possible to quickly build, deploy, operate, and maintain as well as expand services. Based on Docker technology, this research compares and chooses from various authority control models and finally decides to take the role authority management model as the infrastructure. It designs the role authority control model based on cloud computing and Docker technology in combination with the Task Controller Function, the Project Controller Function, and the User Controller Function and realizes this model by adopting the MongoDB database combined with HTML/CSS/Javascript syntax and the Boor Strap framework. After the test, it is found that the Docker technology based role authority management system has satisfactory test performance consistent with expected outputs as well as strong robustness, which can meet the requirements of different objects and subjects.

In the era of data nowadays, role authority management is no longer limited by region or system but is gradually applied to all kinds of systems and has become one of their important components. Role authority management is applied in various systems with the purpose that the module can control and manage the system authorities [1–5]. System management authorities, if not rationally and effectively controlled and managed, will bring about system bugs, information leakages, and other problems that cause serious loss of users [6–10]. Therefore, in order to avoid such problems, it is of great necessity to strengthen the control and management of system authorities. This paper studies the design and realization of the role authority management system based on the cloud computing platform and obtains the implementation interface of the role authority management system based on the cloud computing platform, which enhance the robustness of role authority management system design and development and lower risks.

## 1. Research Status

With the development of mobile Internet technology and the continuous popularization of the network, various kinds of Web software come out in the sight of the public one after another. As the most commonly used function in Web systems, user role authority management is constantly studied.

Research on user role authority in software systems began in 1996 when Professor Ravi Sandhu et al. proposed the RBAC (Role-Based Access Control) model [11], where the concept of user role is introduced for the first time. At present, there are discretionary access control models, mandatory access control models, and role-based access control models. As cloud computing technology and distributed storage technology develop rapidly with the advent of the big data era, data security has become the focus of people's attention, and the access control of cloud systems

has become an urgent problem to be solved. In order to improve the security of traditional access control technology, constraint research based on role access control has become a research hotspot nowadays.

In 1997, Professor Ravi Sandhu proposed the ARBAC97 model (distributed RBAC management model) [12]. On the basis of these two models, the ARBAC99 model and the ARBAC02 model [13] improved the management function of RBAC. After the RBAC96 model, another famous RBAC model is the RBAC2001 model [14] proposed by the American National Institute of Standards and Technology. Up to now, RBAC is still a topic concerned and studied by relevant scholars. According to the relevant research of Mohsen Saffarian et al., they organically integrated DAC and RBAC model to construct the OB-RBAC model. This model can be more applicable to the settings of organizations with higher flexibility or those with lack of standardization. However, when authorities and users increase gradually, this authorization management method will become highly complicated [15, 16].

In China, researchers mainly put forward improvements to the RBAC model in specific practical studies. In order to strengthen the further management of user roles in the system, Fan et al. proposed a role management method based on the RBAC model in 2012 [17]. Zhang et al. proposed a dynamic role access control model based on attribute value variation in 2016 by setting attributes for roles [18], thus realizing the conversion between dynamic roles and static roles. On the basis of retaining the easy authorization and hierarchical management of roles of the RBAC model and in view of the SaaS (Software-as-a-Service) [19] multitenant mode, Xiong et al. proposed a multitenant, multidomain, and secure interoperation access control method based on the role-level tree [20]. In view of the limitations and complementarity of RBAC and ABAC (Attribute-Based Access Control) in large-scale environments, Zhou and Ren proposed an access control model combining attributes and roles in 2018. On this basis, the existing models are divided into three types: dynamic role models, attribute-centered models, and role-centered models [21].

Through the analysis of research status, problems such as low execution efficiency, unbalanced distribution of resources, and complex deployment process are found in the studies abovementioned. Docker provides great convenience to developers with not only standard container specification technology and image specification but also a series of user-friendly container management platforms. It quickly takes up a large market share and many systems have been deployed on Docker. Therefore, this paper employs the operation layer to realize the interaction of front page data, makes advantage of the related methods in the Dao layer to process background data, takes Docker technology as the basis for designing the technical framework, and combines the Task Controller Function, the Project Controller Function, and the User Controller Function to meet the requirements of different objects and subjects. The research idea and the model and framework designed also become the innovation point of this paper.

## 2. Overview of Docker Technology

In cloud computing, PaaS (Platform as a Service) is a service mode that can provide users with software platforms and development environments [22–25]. Docker is an open-source project realized on the basis of Golang and the application of Docker can accelerate the realization of PaaS mode [26]. In software development, we usually encounter software reuse failures caused by development environment difference, while Docker can effectively solve such problems. Whether during the development stage or test stage of software, Docker can create the same environment. In other words, it intervenes in the container configuration to guarantee the consistency of the dependence of all configurations within the container and to enable that the development, test, and release of software all occur in the same environment. Docker has three advantages. First, it can use the image repository. Second, it can conduct continuous deployment test. Third, it possesses high level of resource utilization. In this paper, the role-based authority management system is deployed on the Docker platform. After the server is started and the images are obtained, the installation of operating systems, repositories, and application services will be completed, which greatly shortens the system development process. Just because Docker does not need to start the slave operating system, enough disk space is spared, abundant system resources are saved, and the utilization rate of system resources is improved. It can also be considered that the virtualization of virtual machines is manifested in hardware level and the virtualization of Docker is manifested in operating system level. The architecture of Docker is as shown in Figure 1.

It can be seen in Figure 1 that the Server Docker is the kernel daemon which can be deployed not only locally but also in remote servers. The communication bridge between Server and Client is Rest APL. As the Client, Client Docker CLI can, besides providing corresponding interfaces for users, manage containers and images. After images are packaged, users can create containers by virtue of images and then operate them, during which Server can be invoked to control disk resources. Compose in Docker is employed to conduct rapid arrangement of container clusters before the operation of multiple containers. Practically, it is hard to provide sound operating conditions for the system by relying on only one single container, and the database and server need to be started simultaneously to guarantee the sound system operation. Besides, service and project are two important concepts in Compose. The first means that it can define the services required by the application in terms of the name, network, configuration environment, dependence condition, etc. Service is conducted in the granularity of container and each container carries one service. The second refers to the whole project that needs to be implemented by users. The content (container deployment, container management, etc.) disclosed in YMAL file can be deemed as a project and Compose will usually manage the project without extra interference.



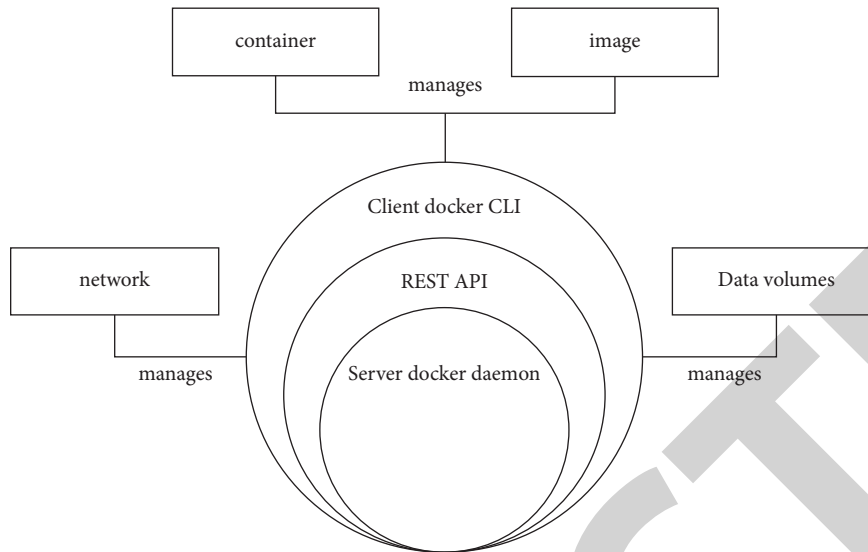


FIGURE 1: Architecture of Docker.

### 3. Model Building and Design of Background Authority Functions

**3.1. Basic Model Building.** Build the authority management system model on the basis of the role-based access control technology. Due to the large scale of the system, we can not conduct authority setting and role division for every user in traditional ways. The concept of user group should be introduced and users with the same authorities and characteristics should be divided into the same user group without changing the many-to-many relationship in role-based access control. One user can be divided into multiple user groups and one user group relates to multiple users. This can not only reduce the heavy and complicated workload but also facilitate subsequent management. Take the task management system as an example. Users can be divided into two user groups: first, the project participant user group, and second, the project administrator user group. Different user groups have different authorities. For example, roles in the administrator user group have authorities to modify and delete the project while roles in the participant user group only have authorities to preview and edit the tasks but cannot modify or delete the project. See Figure 2.

**3.2. Comparison and Option of Authority Control Models.** In user information management systems, authority control is one necessary link. The so-called authority means that users need to access authorized content in accordance with the security rules set by the system. Authority control can effectively identify user identity and avoid illegal invasion; therefore, authority control has been widely promoted in numerous user login systems. Current authority control technologies can fall into two categories: the system level security management, such as operating system level security management and database level authority management, etc., and the application-level security management which generally is closely related to specific system

requirements. From the perspective of implementation, authority control mainly employs three types of models: (1) the discretionary access control model; (2) the mandatory access control model; and (3) the role-based access control model.

**3.2.1. Discretionary Access Control.** It means that the subject with control power can authorize the access right of the object to other subjects [27]. In such access management model, one user can have different authorities to different resource objects while different users can have different authorities to one same resource object; one user can authorize his authority to other users without any limitation. In the discretionary access control model, users can formulate corresponding protection strategies for the resource object to be protected according to system requirements. Discretionary access control has advantages such as flexible authority distribution, simple and easy-to-use models, and strong expansibility; however, the discretionary pattern of authorization leads to low security level of the whole system while the high complexity of discretionary authorization is also a hard nut to crack.

**3.2.2. Mandatory Access Control.** Its main function is to prevent virus attacks and ensure the confidentiality and security of information [28]. The access control method is to mark the security of the elements in the system, that is, to give each subject (user, process) and object (file, data) a security level and restrict the access behavior of the subject to the object by comparing their different security levels. In the mandatory access control model, access control is executed mandatorily by the system and users do not have authority to change their security level. The most typical mandatory access control model is the Bell-LaPadula model, the relevant access control rules of which are as shown in Table 1.

It can be seen in Table 1 that there are four security levels for objects and subjects which are A (top-secret), B (secret),

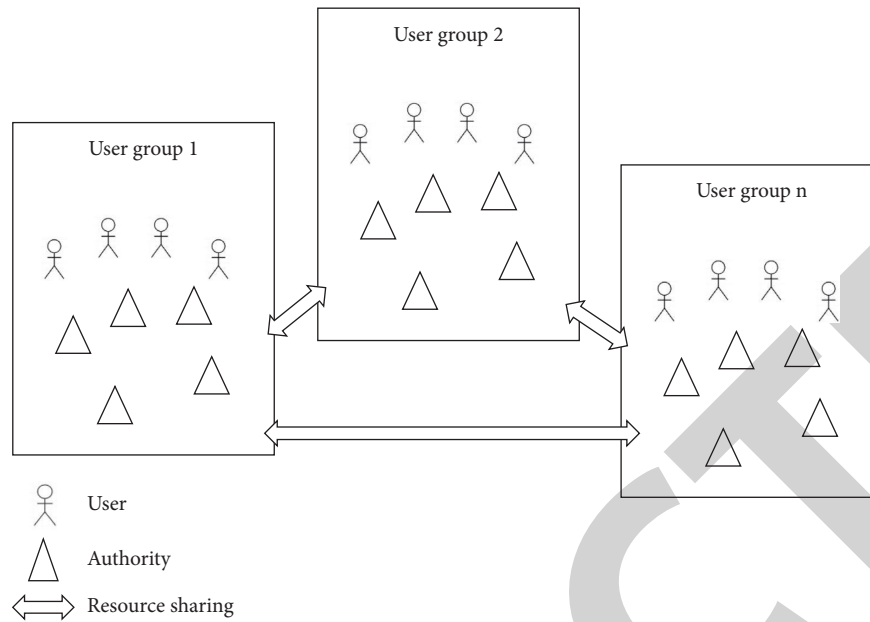


FIGURE 2: Basic model architecture of the authority management system.

TABLE 1: Decision rules.

		Object security level			
		A	B	C	D
Subject security level	A	I/Q	I	I	I
	B	Q	I/Q	I	I
	C	Q	Q	I/Q	I
	D	Q	Q	Q	I/Q

C (confidential), and D (unclassified) successively from top to bottom. Read and write operations are represented by I and Q in Table 1, respectively. The access control rules are as follows: (1) when the security level of the subject is higher than that of the object, the subject can execute the read operation to the object; (2) when the security level of the subject is lower than that of the object, the subject can execute the write operation to the object. Such feature is usually called “Read Down, Write Up.” Such access rules can effectively prevent subjects of lower security level from accessing resources of higher security level, which guarantees information confidentiality.

**3.2.3. Role-Based Access Control.** Such model successfully integrates the concept of role into users and authorities [29–31]. Users can play their required part of roles in a reasonable scope and all roles have their authorities and then work accordingly. The core structure of this model is as shown in Figure 3.

In this model, users and authorities are not completely corresponding for one user can play the part of multiple roles and one role can have multiple users on the contrary. As well, the relation between roles and authorities is not single but has many-to-many characteristics. The relation of the three is as shown in Figure 4.

Role-based access control has advantages such as increasing the flexibility of the system and facilitating the distributed management of the administrator. Such model can effectively solve the problems of system management complexity and security defects in access control and meanwhile design authority schemes that satisfy the requirements of roles by virtue of database interfaces. From this, such model can be regarded as a user + role + authority model.

In summary, comparing with the discretionary access control model and the mandatory access control model, the robe-based authority management model has remarkable advantages in terms of flexibility, security, and management modes; therefore, the role authority management model is selected as the system authority control model in this paper.

**3.3. Design of Background Authority Functions.** In the background architecture, the operation layer can facilitate the interaction of front page data and the related methods in the Dao layer can help with the processing of background data. Therefore, three types of background authority control functions are set here, namely, the Task Controller Function, the Project Controller Function, and the User Controller Function, so as to meet the requirements of different objects and subjects.

## 4. Design and Realization of Cloud Computing Based Role Authority Management Systems

### 4.1. Design of Cloud Computing Based Role Authority Management Systems

**4.1.1. Design of Database.** After the authority control scheme is implemented, the background database side is needed to conduct tests. For MongoDB database belongs to

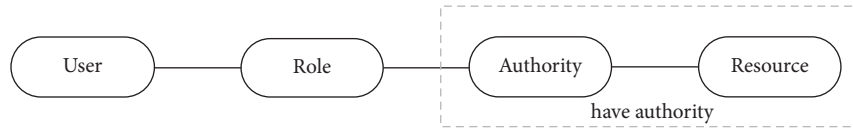


FIGURE 3: Model core structure.

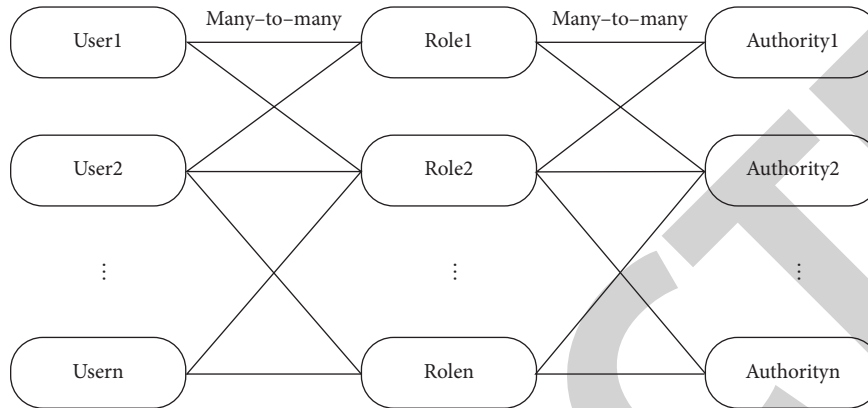


FIGURE 4: Relation diagram of users, roles, and authorities.

distributed file storage NoSQL database, interactive orders can be employed for the goal of operating and accessing the database; meanwhile the storage structure is flexible and changeable, and therefore Spring Boot is utilized to facilitate the design and realization of MongoDB database. Besides, to guarantee the effectiveness of authority control, the view mode is utilized to measure the authority of user groups. Specifically speaking, directly see the authorities possessed by roles in every user group in the form of views and then operate in the page accordingly.

**4.1.2. Design of Database Background Architecture.** The database background architecture is mainly composed of the Dao layer, the entity layer, and the operation layer. The entity layer here includes three entity classes, namely, the project, task, and user. The Dao layer enables operations such as deleting, adding, modifying, and querying of the database. The operation layer contains three related files, namely, the project, task, and user files. This layer can realize the front page data interaction; besides, it can also reasonably invoke the background data. And corresponding methods can be quickly found in the project class.

**4.1.3. Front Page Design and Parameter Passing.** HTML/CSS/Javascript grammars are utilized to realize the development of front pages and the Boor Strap framework is introduced to obtain better front pages. To guarantee the effective interaction between the front-end and rear-end pages, Spring Boot is also utilized here. Besides, in returning JSON objects to the front end, the parameter passing employs the value of Ajax function in JQuery.

**4.2. Realization of Cloud Computing Based Role Authority Management Systems.** For the test system utilizes Docker

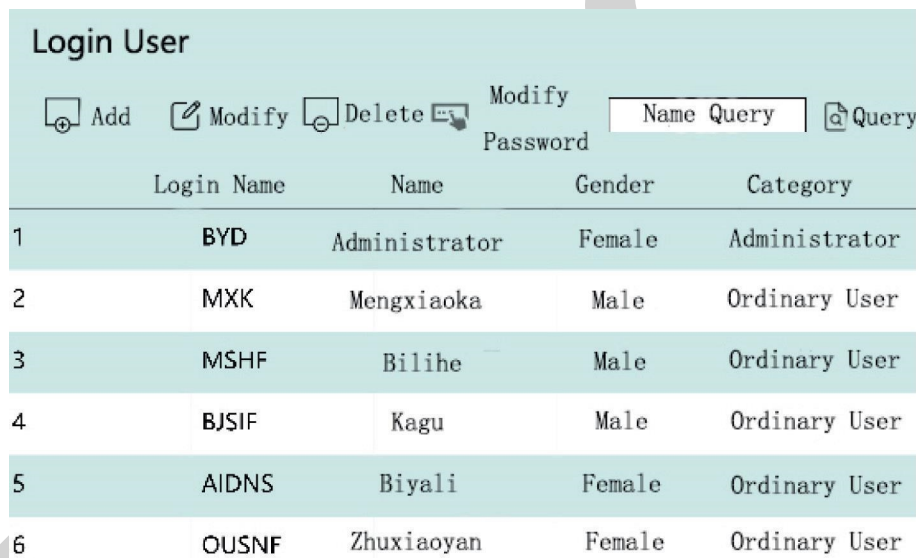
during the deployment, the environment and dependence required by the project should be stored. After the image construction is completed, the whole set of service can be obtained by directly operating images. Everyone who wants to test the project effect can test on his own simply by laying Docker environment in his host to build images and thus to reduce the tedious test process for the test personnel.

**4.2.1. Configuration Environment.** The hardware environment package of the cloud computing platform has multiple physical hosts matched with X86-64 chips. All the physical hosts carry Docker environments, chose Linux as the bottom layer operating system, and support 64-bit environment. Under the Ubuntu system, there are three steps in the configuration environment flow. First, examine the preparatory condition “uname-a” of the machine, upgrade the system kernel when it is found that the system kernel cannot meet the requirements of the design, and execute relevant installation operations only after the upgrading is completed. Meanwhile, to guarantee the normal installation of Curl order, add APT repository for assistance. Second, add GPG keys, renew the APT source, and then complete the installation of Docker software packages. Third, test the installation effect of Docker and execute “sudo docker info” order if the installation is determined to be normal. When no container or image is found in the order, the Docker environment is not completely installed, and we need to wait or reinstall it; otherwise, the installation of Docker environment is completed.

**4.2.2. Deployment File Writing.** Docker-compose.yml and Dockerfile are the two core files necessary for the system deployment. In docker-compose.yml file, two types of images need to be written-in. The first is the images of the

TABLE 2: System test results.

Item	Input/action	Expected output	Actual test result
Authority management	Users clicking the “Authority Management” button	Entering the role authority management interface	Test result consistent with the expected output
	Finding relevant content in the Find bar as needed	Displaying information found	Test result consistent with the expected output
	Clicking the “Modify” button	Displaying “Modify Authority”	Test result consistent with the expected output
Role management	Users clicking the “Role Management” button	Entering the role management interface	Test result consistent with the expected output
	Clicking the “Add” button	Displaying the user addition interface	Test result consistent with the expected output
	Clicking the “Modify” button	Displaying the user role modification interface	Test result consistent with the expected output
Node management	Clicking the “Query” button	Displaying all node definition information	Test result consistent with the expected output
	Clicking the “Delete” button	Displaying corresponding node to delete interface	Test result consistent with the expected output



The screenshot shows a web interface titled "Login User". At the top, there are several action buttons: "Add", "Modify", "Delete", "Modify", "Name Query", and "Query". Below these buttons is a table with the following columns: "Login Name", "Name", "Gender", and "Category". The table contains six rows of user data:

	Login Name	Name	Gender	Category
1	BYD	Administrator	Female	Administrator
2	MXK	Mengxiaoka	Male	Ordinary User
3	MSHF	Bilihe	Male	Ordinary User
4	BJSIF	Kagu	Male	Ordinary User
5	AIDNS	Biyali	Female	Ordinary User
6	OUSNF	Zhuxiaoyan	Female	Ordinary User

FIGURE 5: A part of the implementation interface of the cloud computing platform based role authority management system.

Spring Boot project. Spring Boot is a new framework developed by the Pivotal team to ease the difficulty of developing and building new Spring. Spring Boot is configured in a specific way rather than in traditional ways, avoiding manual configuration by developers. With this advantage, Spring Boot has been greatly promoted in the field of application development and has become the first choice of developers for system development. Spring Boot not only inherits the advantages of the Spring framework, but also optimizes the configuration, which further eases the difficulty of developing and building Spring applications. In addition, after the integration of a large number of frameworks, the problems of dependent package version conflict and application uncertainty are solved. The second is the images of MongoDB, the customization process of which is as follows: first, customize an nginx image and store the image in /usr/share/nginx/html/index.html file; second, input the From order and the RUN order; and third, execute

the building action under the storage directory of Dockerfile. After the images are successfully built, it is necessary to set the context path and construct the image instructions, such as constructing the parameter instructions, defining the anonymous data volume instructions, constructing the port declaration instructions, and establishing the working directory instructions. The system can only operate normally after the two types of images are written-in, while in Dockerfile, the environment and dependence required by the project should be defined to realize the independent compilation of Docker.

**4.2.3. Project Deployment.** Invoke and execute the “docker-compose up-d” order in the project, and then relevant images are obtained and the system is operated. The operation state of the containers can be directly checked through the “docker logs” order other than conducting

secondary examination of the MongoDB containers because the information of MongoDB is integrated into the images of the Spring Boot project. When MongoDB fails to start, the start of the project will terminate. The specific working process of the project is as follows: first, monitor the user's URL request; second, analyze the user request and filter the URL content; third, complete the selection of back-end processing functions according to different parameters and request contents; fourth, call the back-end processing function and get response; and fifth, transfer the content that needs to be returned to the user and present it on the front page.

## 5. System Test

After the deployment is completed, the system is tested to verify the sound interaction between the cloud computing platform and the role authority management system. The interaction service is carried out at the entrance of the user access system, which enables users to visually learn about the real-time status of the project and the requests of users can be processed accordingly. Details are as shown in Table 2.

It can be seen that the cloud computing platform based role authority management system has satisfactory test results consistent with expected outputs as well as strong robustness. A part of the implementation interface of the cloud platform based role authority management system test is presented as shown in Figure 5.

It can be seen that the cloud computing platform based role authority management system has complete functions and can effectively disclose member information. The cloud computing platform and role authority management system have good interaction, which confirms the feasibility of the design of this paper.

## 6. Conclusion

This paper first gives an overview of the research in this field, analyzes authority control models, Docker, and other related technologies, determines the role authority control model based on the basic framework of Docker, and designs the role authority management system on this basis. Finally, the cloud computing platform is deployed, its interaction with the role authority management system is completed, and the core authority function is realized. Through the test, the function of the cloud computing platform based role authority management system is fully verified, and the feasibility of the model design is confirmed. With the continuous research and development as well as optimization of the model, the basic framework can completely replace traditional role authority management modes and then promote the development of role authority management.

Docker is a lightweight virtual technology with advantages of shorter deployment time and higher resource utilization rate than virtual machines. The deployment of the project on Docker enables the system to realize the purpose of multiple operations at one time, which greatly shortens development and testing time, improves work efficiency, and is highly practical. However, this technology also has some

problems such as poor isolation performance and serious waste of storage resources, which affects the performance of the design architecture in this paper and indicates the direction for future research.

## Data Availability

All the data contained in this study can be obtained by contacting the corresponding author. Readers can also inquire part of the original data and the results of data processing in this paper.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported by 2021 Hebei Human Resources and Social Security Research Cooperation Projects (JRSHZ-2021-01080).

## References

- [1] S. Q. Deng, "Research on the design of goal-oriented rights management system," *Industrial Design*, vol. 6, pp. 67–69, 2021.
- [2] W. L. Li, X. J. Jin, and H. Zhao, "Rights management system based on resource control design method," *Computer Knowledge And Technology*, vol. 3, pp. 44–45, 2021.
- [3] D. Q. He, H. Hu, and J. C. Yan, "General rights management system based on rbac," *Computer Knowledge And Technology*, vol. 33, pp. 97–102, 2020.
- [4] P. Zhang, S. I. Li, and Y. M. Liu, "Research on resource scheduling in edge heterogeneous computing system powered by block chain," *Journal of Communications*, vol. 41, no. 10, pp. 1–14, 2020.
- [5] P. Y. Fu, J. Wu, and X. F. Zhang, "And steady-state genetic algorithm based on edge computing aiot resource scheduling research," *Journal Of Xiangtan University (Natural Science Edition)*, vol. 5, pp. 71–83, 2020.
- [6] M. X. Li, C. Cao, Y. Tang, and Q. Liu, "Research on edge resource scheduling solution for computing power network," *Frontiers Of Data And Computing*, vol. 2, no. 4, pp. 80–91, 2020.
- [7] A. Z. Fu and X. R. Wang, "Research and implementation of authority management for storage management system based on Apache shiro," *Computer Programming Skills And Maintenance*, vol. 3, pp. 99–102, 2020.
- [8] Q. C. Zhou, Y. T. Wang, J. Zhao, and X. L. Xiong, "Design and implementation of fine-grained permission management system based on rbac model," *Network Security Technology And Application*, vol. 10, pp. 38–41, 2019.
- [9] Z. H. Qu, R. Ye, and G. H. Chen, "Research progress of resource optimization technology for edge computing," *Big Data*, vol. 5, no. 2, pp. 17–33, 2019.
- [10] Z. W. Liang, "Design and application of general permission management system for large-scale information system," *Information Technology And Informatization*, vol. 9, pp. 38–40, 2018.
- [11] R. S. Sandhu, E. J. Coyne, H. L. Feinstein, and C. E. Youman, "Role-based access control models," *Computer*, vol. 29, no. 2, pp. 38–47, 1996.

## *Retraction*

# **Retracted: Digital Protection and Development of Intangible Cultural Heritage Relying on High-Performance Computing**

## **Computational Intelligence and Neuroscience**

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

## **References**

- [1] H. Ding, "Digital Protection and Development of Intangible Cultural Heritage Relying on High-Performance Computing," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4955380, 10 pages, 2022.



## Research Article

# Digital Protection and Development of Intangible Cultural Heritage Relying on High-Performance Computing

**Hong Ding** 

*Chengdu Normal University, Chengdu, Sichuan 611130, China*

Correspondence should be addressed to Hong Ding; 561044@cdnu.edu.cn

Received 28 January 2022; Accepted 17 March 2022; Published 13 April 2022

Academic Editor: Daqing Gong

Copyright © 2022 Hong Ding. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The protection of intangible cultural heritage has gradually attracted people's attention. In today's digital information age, the use of digital technology in the protection of intangible cultural heritage has become the focus of research in the field of intangible cultural heritage protection. For the protection of intangible cultural heritage to become the key goal of the continuous inheritance of China's traditional culture, this paper fully analyzes the problems existing in the inheritance and protection of intangible cultural heritage, and applies the high-performance computing method to the digital protection and development of intangible cultural heritage, which provides a strong basis for the protection of intangible cultural heritage. The accurate identification and protection of intangible cultural heritage can be achieved by using high-performance computing methods. In addition, the digital protection and development system of intangible cultural heritage is also designed. Finally, the example results show that the digital protection and development system of intangible cultural heritage proposed in this paper can effectively protect intangible cultural heritage and promote the continuous development of domestic local economy. At the same time, it can also strengthen the public awareness of the protection of intangible cultural heritage and effectively promote the inheritance and development of the protection of intangible cultural heritage.

## 1. Introduction

China's intangible cultural heritage has rich cultural connotation and cultural heritage value. In the process of continuous excavation, it can reproduce historical events and development tracks. Meanwhile, intangible cultural heritage also carries Chinese traditional cultural beliefs and inherits historical culture in a variety of ways, carries the historical culture with diversity and flexible development as a long-term development system of Chinese civilization. It can reflect the value of national tradition, cultural value and aesthetic value. The protection and inheritance of intangible cultural heritage are inherent spiritual needs of human beings [1–3]. However, the aesthetic status of intangible cultural heritage deserves further consideration, which is confirmed in the In terms of the continuous development of society, the continuous economic growth, the inheritance of traditional culture and the awareness of cultural values, and can be carried out in the protection and development of traditional intangible cultural relics, and the electronic

archives of intangible cultural heritage is established and other multi-channel protection methods are utilized to protect and develop the traditional intangible cultural heritage [4–6]. With the continuous progress of society, in the process of development and protection of intangible cultural heritage, digital protection will become the goal of in-depth research and exploration by Chinese and foreign scholars. The digital technology is used to electronically archive intangible cultural heritage, not only effectively processing massive data, be conducive to data storage and query, but also effectively avoid the loss and damage of data and information caused by natural and man-made disasters [7, 8]. The data mining technology can be used to deeply analyse the stored intangible cultural heritage data, provide valuable information data for visitors or users, and then complete the exchange and application of culture around the world. the high-performance computing and identification can be performed for the intangible cultural heritage as one of the most popular types of visual aggregation, which has become the focus of Chinese scholars' research [9–11].



However, how to obtain specific content sources from the initial intangible cultural heritage data that lacks the definition of intangible cultural heritage content has become a huge challenge for current high-performance computing of intangible cultural heritage, because intangible cultural heritage signals follows the way of time -sorting, which can use the high-performance computing in accordance with its stealth. At present, other classification methods are relatively simple, and the characteristics of intangible cultural heritage obtained are not accurate.

This paper proposes the digital protection and development of intangible cultural heritage. This method uses high-performance computing methods to extract the characteristics of intangible cultural heritage, calculate the weight of its protection space, integrate the space with high similarity in meaning, and build an intangible cultural heritage protection and development model. Intangible cultural heritage preserves cultural integrity through digitization, which is different from traditional cultural expressions and forms, because it is not a physical existence, but an information exchange across regions and countries. It covers scientific culture, social environment, civilized science and other cultures.

## 2. Technologies and Methods

**2.1. High-Performance Computing Methods.** The high-performance computing system can realize the I/O operation of the hard disk as little as possible based on the memory computing technology, and can also perform operation data analysis in the memory. Hadoop's MapReduce calculation engine saves the calculated intermediate value in the memory. In the next calculation, the original I/O request must be stored externally. Since high-performance computing directly stores intermediate values in memory and repeats them, high-performance computing reduces the time of data transmission as a whole compared with previous MapReduce, thereby ensuring the work efficiency of tasks [12, 13].

In view of the accuracy and diversity characteristics of intangible cultural heritage based on high-performance computing algorithms, the comprehensive analysis standard representation of data based on high-performance computing algorithms is defined, including

$$Eval(\pi_p) = \lambda Acu(\pi_p) + (1 - \lambda) Di v(\pi_p). \quad (1)$$

In the formula  $\lambda \in [0, 1]$ , the correctness and diversity of data processing are important degrees in the comprehensive analysis standard.

Formula (1) calculates the probability  $pro(\pi_p)$  that each data basic processing algorithm is selected as the analysis basic processing based on the diversity  $Div(\pi_p)$  of the basic processing of the high-performance computing algorithm. The calculation formula is as follows.

$$pro(\pi_p) = \frac{Div(\pi_p)}{\sum_{p=1}^B Div(\pi_p)}. \quad (2)$$

The purpose of the data processing network module is to make the color and spatial position of each reconstructed data to restore the inherent color and texture of the data. The loss function  $Lip$  of the data processing network module is defined by the formula. The above shown are processing loss for unmasked regions, processing loss for masked regions, perceptual loss, style loss, antiloss, and total difference loss.

$$L_{total}^{inp} = 2L_{valid} + 12L_{hole} + 0.04L_{per} + 100(L_{style}^1 + L_{style}^2) + 100L_{a_{dv}} + 0.3L_{var}. \quad (3)$$

The weight of each loss item is determined by analyzing the results of 50 independent experiments.

Manhattan distance uses the unmasked region of both the processed mode and the real mode to process the loss. In the equation,  $I_{dam}$  is the damage mode,  $M$  is an irregular binary mask (the corresponding processing area in the mask is 0, the others are 1), and  $I_{inp}$  is the processing result mode.  $I_{real}$  actually indicates there is no damage mode. The processing loss function for masked regions is similar to the expression:

$$L_{valid} = \|M \times (I_{inp} - I_{dam})\|_1, \quad (4)$$

$$L_{hole} = \|(1 - M) \times (I_{inp} - I_{dam})\|_1,$$

High-performance computing algorithms dynamically adjust appropriate parameter values based on the identification complexity of different regions of the data, so that the high-performance computing algorithm maintains a constant speed, and the generated metadata is more neatly unified in subsequent processing. In order to reduce the impact of the difference in the shape of the super data on the algorithm, this parameter does not need to be specified. Compare the  $I_{real}$  reduction identification points of intangible cultural heritage.  $h$  is the junction point of the first data, which is located in the data corresponding to the intangible cultural heritage database.  $0 \leq k \leq h(vx'_{2k+1,i}, vy'_{2k+1,i})$ , the following mining operations are performed on all identification points:

$$\begin{cases} dx'_k = vx'_{2k+1,i} - vx_{k,i} \\ dy'_k = vy'_{2k+1,i} - vy_{k,i} \end{cases},$$

$$\begin{bmatrix} vx'_{2k+1,i} \\ vy'_{2k+1,i} \end{bmatrix} = \begin{bmatrix} vx_{k,i} \\ vy_{k,i} \end{bmatrix} + \begin{bmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{bmatrix} \times \begin{bmatrix} dx'_k \\ dy'_k \end{bmatrix}. \quad (5)$$

For each piece of data  $P_i$  and  $\tilde{P}_i$ , calculate the barycentric coordinates after removing the last point:

$$\begin{cases} \overline{vx}'_i = \frac{1}{h'_i - 1} \sum_{k=1}^{h'_i - 1} vx'_k \\ \overline{vy}'_i = \frac{1}{h'_i - 1} \sum_{k=1}^{h'_i - 1} vy'_k \end{cases} \begin{cases} \overline{vx}_i = \frac{1}{h_i - 1} \sum_{k=1}^{h_i - 1} vx_k \\ \overline{vy}_i = \frac{1}{h_i - 1} \sum_{k=1}^{h_i - 1} vy_k \end{cases} \quad (6)$$

In the formula,  $h$  is the length (number of nodes) of the  $i$ -th piece of data after the intangible cultural heritage database is divided; the  $h'_i$  data is embedded in the identified length, and the following two offset values are calculated:

$$\begin{cases} \Delta x_i = \frac{1}{2h_i - 2} \sum_{k=1}^{h_i-1} (vx_{k+1} - vx_k) P_k, \\ \Delta y_i = \frac{1}{2h_i - 2} \sum_{k=1}^{h_i-1} (vy_{k+1} - vy_k) P_k, \end{cases} \quad (7)$$

According to formula (12)~formula (15), calculate the identification point of the vertical and horizontal coordinates of each piece of data respectively.

①  $\Delta x_i \neq 0, \Delta y_i \neq 0$

$$\begin{cases} qx_j = \sum_{\{i|i/c=j\}} \frac{\overline{vx'_i} - \overline{vx_i}}{\Delta x_i} = \sum_{\{i|i/c=j\}} \tilde{b}_i \alpha = \tilde{c} \tilde{b}_j \alpha, \\ qy_j = \sum_{\{i|i/c=j\}} \frac{\overline{vy'_i} - \overline{vy_i}}{\Delta y_i} = \sum_{\{i|i/c=j\}} \tilde{b}_i \alpha = \tilde{c} \tilde{b}_j \alpha. \end{cases} \quad (8)$$

②  $\Delta x_i = 0, \Delta y_i \neq 0$

$$\begin{cases} qx_j = \sum_{\{i|i/c=j\}} \sum_{k=1}^{h_i-2} \frac{\overline{vx_{k,i}} - vx_{k,i}}{(vx_{k+1,i} - vx_{k,i}) P_k} = \sum_{\{i|i/c=j\}} \tilde{b}_i \alpha = \tilde{c} \tilde{b}_j \alpha, \\ qy_j = \sum_{\{i|i/c=j\}} \frac{\overline{vy'_i} - \overline{vy_i}}{\Delta y_i} = \sum_{\{i|i/c=j\}} \tilde{b}_i \alpha = \tilde{c} \tilde{b}_j \alpha. \end{cases} \quad (9)$$

③  $\Delta x_i \neq 0, \Delta y_i = 0$

$$\begin{cases} qx_j = \sum_{\{i|i/c=j\}} \frac{\overline{vx'_i} - \overline{vx_i}}{\Delta x_i} = \sum_{\{i|i/c=j\}} \tilde{b}_i \alpha = \tilde{c} \tilde{b}_j \alpha, \\ qy_j = \sum_{\{i|i/c=j\}} \sum_{k=1}^{h_i-2} \frac{\overline{vy_{k,i}} - vy_{k,i}}{(vy_{k+1,i} - vy_{k,i}) \circ P_k} = \sum_{\{i|i/c=j\}} \tilde{b}_i \alpha = \tilde{c} \tilde{b}_j \alpha. \end{cases} \quad (10)$$

④  $\Delta x_i = 0, \Delta y_i = 0$

$$\begin{cases} qx_j = \sum_{\{i|i/c=j\}} \sum_{k=1}^{h_i-2} \frac{\overline{vx_{k,i}} - vx_{k,i}}{(vx_{k+1,i} - vx_{k,i}) P_k} = \sum_{\{i|i/c=j\}} \tilde{b}_i \alpha = \tilde{c} \tilde{b}_j \alpha, \\ qy_j = \sum_{\{i|i/c=j\}} \sum_{k=1}^{h_i-2} \frac{\overline{vy_{k,i}} - vy_{k,i}}{(vy_{k+1,i} - vy_{k,i}) \circ P_k} = \sum_{\{i|i/c=j\}} \tilde{b}_i \alpha = \tilde{c} \tilde{b}_j \alpha. \end{cases} \quad (11)$$

Calculate the cultural information value for each piece of data:

$$\tilde{m}_j = \begin{cases} 1, \frac{qx_j + qy_j}{2} > 1, \\ 0, \frac{qx_j + qy_j}{2} < 1. \end{cases} \quad (12)$$

After extracting the cultural information, check the authenticity of the identification. This process uses the correlation coefficient  $cor(m, \tilde{m})$  to detect the similarity between the original recognition and the extraction recognition:

$$cor(m, \tilde{m}) = \frac{\sum_{i=0}^{n-1} (m_i \tilde{m}_i)}{\sqrt{\sum_{i=0}^{n-1} m_i^2} \sqrt{\sum_{i=0}^{n-1} \tilde{m}_i^2}} \quad (13)$$

In the formula,  $\tilde{m}$  is the cultural information extracted from the identified graphics;  $cor(m, \tilde{m})$  is the correlation coefficient of the cultural information  $m$  and  $\tilde{m}$ . When high-performance computing optimizes the BP algorithm, the roulette method is usually chosen. If the fitness of the  $i$ -th individual is  $f_i$ , the probability  $P_{si}$  of the  $i$ -th individual being left behind is

$$P_{si} = \frac{f_i}{\sum_{j=1}^n f_j}. \quad (14)$$

In this paper, as shown in (9), an overall crossover scheme is chosen.

$$\alpha = \begin{cases} \frac{f_{\max} - f_i}{f_{\max} - f_{avg}}, f_i \geq f_{avg}, \\ 0.35, f_i < f_{avg}. \end{cases} \quad (15)$$

In the formula,  $\alpha$  represents the gene-encoded combination coefficient,  $f_i$  represents the fitness value of the  $i$ -th individual in the parent population,  $f_{\max}$  and  $f_{avg}$  represent the maximum and average fitness values of the individuals in the parent population, respectively.

The final high-performance computing process is shown in Figure 1.

Effectively combining the compatibility of high-performance computing with more distributed systems, the main operating modes of the system can be divided into single-point mode and dispersion mode. This mode mainly uses YARN, Mesos or EC2 for resource allocation and scheduling. RDD is computed by using high-performance computing, and the memory with prominent operation characteristics can effectively improve the calculation efficiency of the model. RDD has the data characteristics of compatibility and parallelism, and can interact with the memory and external memory. If the storage cannot meet the requirements of a single RDD, the RDD will be transferred to the external memory. The digital protection and development algorithm can obtain better performance in the process of data mining. This algorithm can be used for training data to reduce the accuracy and performance. In order to solve this problem, the data mining method is used to directly determine the

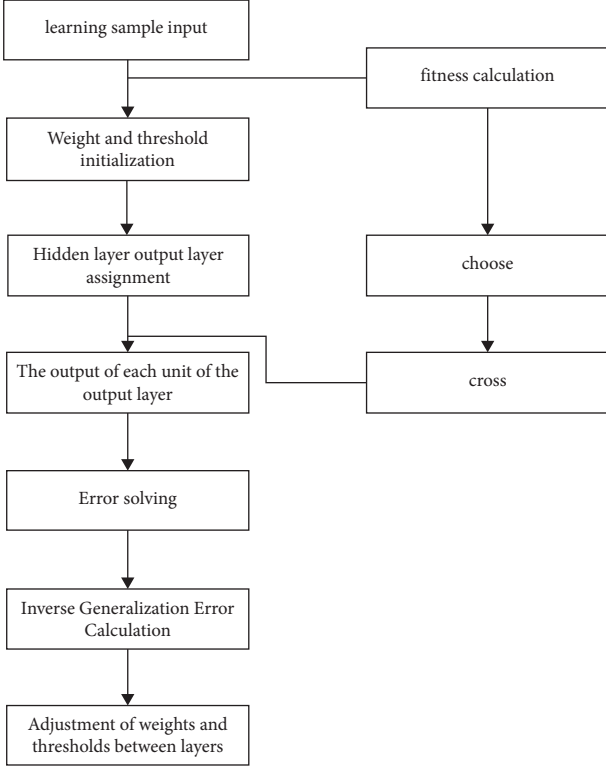


FIGURE 1: High-performance computing flow chart.

first category as the final data. As a result, the tie situation that cannot be solved by the average value cannot be fundamentally solved, so the drawbacks of the model can be effectively avoided. The classification performance of each classifier is protected using the Berius equation. The formulas that can be used are:

$$\text{weight}(i) = 1 - \frac{1/\text{con}(i)}{\sum_{j=1}^T 1/\text{con}(j)} - \frac{N-2}{N}. \quad (16)$$

In formula (1),  $N$  represents the total number of classifiers,  $\text{con}(k)$  represents the posterior probability of the classification result of the  $k$ th classifier in the Bagging algorithm, and  $\text{weight}(i)$  represents the weight of the  $i$ -th classifier.

Two properties of the formula:

- (1) The higher the classification accuracy rate, the higher the weight of the high-performance computing
- (2) Conform to the formula  $\sum_{j=1}^T \text{weight}(i) = 1$ , that is, all high-performance computing can meet the normalization conditions

The Bagging algorithm can be analyzed through out-off-bag data. The F1 value of out-off-bag is smaller than the Bayes formula according to the calculated verification probability. Therefore, in order to reduce the amount of calculation and give up the post-probability of the classification result, the F1 of Out-of-bag is used to judge the correctness of the classification. You can obtain formula (12) by modifying formula (11):

$$\text{weight}(i) = \frac{2}{N} - \frac{1/\text{oob}F_1}{\sum_{j=1}^T 1/\text{oob}F_1(j)}. \quad (17)$$

In formula (2),  $N$  represents the total number of classifiers,  $\text{oob}F_1(k)$  represents the F1 value of the out-of-bag of the  $k$ th classifier, and  $\text{weight}(i)$  represents the weight of the  $i$ -th classifier.

From the perspective of industrial connections, the collection of enterprises generally appears as a vertical chain connection, upstream and downstream enterprises are effectively linked through production links, and the cooperation and competition relationship is the center between enterprises. The SDA method assumes that  $n$  sample sets  $X$  are processed.

$$X = \{x_1, x_2, \dots, x_n\}. \quad (18)$$

The index used is  $p$ , and fuzzy clustering is performed according to the classification:

$$x_k = \{x_{k1}, x_{k2}, \dots, x_{kp}\}, x_{kj} = [a_{kj}, b_{kj}], 1 \leq k \leq n, 1 \leq j \leq p. \quad (19)$$

The relative membership degree of the matrix of the model is represented by  $U$ :

$$U = \{u_{ik}\}, (i = 1, 2, \dots, c; k = 1, 2, \dots, n). \quad (20)$$

$u_{ik}$  indicates the membership degree of sample point  $k$  to category  $i$ , and satisfies the following conditions.

$$\begin{cases} \sum_{i=1}^c u_{ik} = 1, \forall k, \\ 0 \leq u_{ik} \leq 1, \forall k, i. \end{cases} \quad (21)$$

The cluster center of the  $i$ -th category can be represented by  $g_i$ :

$$g_i = (g_{i1}, g_{i2}, \dots, g_{ip}), g_{ij} = [\alpha_{ij}, \beta_{ij}], 1 \leq i \leq c, 1 \leq j \leq p. \quad (22)$$

Here, the weight of the cluster is expressed by importing the adaptive parameter  $\lambda$ :

$$\lambda_k^m = (\lambda_{k1}^m, \lambda_{k2}^m, \dots, \lambda_{kp}^m). \quad (23)$$

The expression of comprehensive weight is

$$W = \sum_{i=1}^c \sum_{j=1}^n (u_{ik})^2 \Phi(x_k, g_i) = \sum_{i=1}^c \sum_{j=1}^n (u_{ik})^2 \cdot \sum_{j=1}^p \left[ \lambda_k^m \left( \frac{a_{kj} + b_{kj}}{2} - \frac{\alpha_{kj} + \beta_{kj}}{2} \right)^2 \right], \quad (24)$$

which satisfies

$$\begin{cases} \lambda_{ij}^m \geq 0, \\ \prod_{j=1}^p \lambda_{ij}^m = 1. \end{cases} \quad (25)$$

The Lagrangian method is used to derive the aggregation function to obtain

$$\begin{aligned}
\alpha_{ij} &= \frac{\sum_{i=1}^n (u_{ik})^2 a_{kj}}{\sum_{i=1}^n (u_{ik})^2}, \beta_{ij} = \frac{\sum_{i=1}^n (u_{ik})^2 b_{kj}}{\sum_{i=1}^n (u_{ik})^2}, \\
\lambda_{ij}^m &= \frac{\prod_{h=1}^p \left[ \sum_{k=1}^p (u_{ik})^2 (a_{kh} + b_{kh}/2 - \alpha_{ih} + \beta_{ih}/2)^2 \right]^{1/p}}{\sum_{k=1}^p (u_{ik})^2 (a_{kj} + b_{kj}/2 - \alpha_{ij} + \beta_{ij}/2)^2}, \\
u_{ik} &= \left[ \frac{\sum_{h=1}^c \sum_{j=1}^p \left[ \lambda_{ij}^m (a_{kh} + b_{kh}/2 - \alpha_{ih} + \beta_{ih}/2)^2 \right]}{\sum_{h=1}^c \sum_{j=1}^p \left[ \lambda_{hj}^m (a_{kh} + b_{kh}/2 + \alpha_{ih} + \beta_{ih}/2)^2 \right]} \right].
\end{aligned} \tag{26}$$

Generally speaking, the effect of enterprise concentration mainly comes from various factors such as natural resources, geographical environment and process technology. The initial integration mainly reflects the production and operation cooperation between different enterprises. Geographical location integration among enterprises is essentially based on competitive and cooperative relations, and the clustered enterprise matching degree is used for integrating internal resources of the enterprise, reducing trade costs as much as possible, and maximizing profits at the same time.

High-performance computing can build the most serialized time series models and use them on a large scale in the process of digital protection and development of intangible cultural heritage. The N states existing in the model can be set to  $S = \{S_1, S_2, \dots, S_n\}$ , then the state at time  $t$  is represented by  $q_t$ . Using  $A = \{a_{ij}\}$  in the transition matrix between different states, the following expression can be obtained:

$$\alpha_{ij}(k) = P[q_{t+1} = S_j | q_t = S_i], 1 \leq i, j \leq N. \tag{27}$$

For some samples, any state can reach other states in one transition; other samples, In other samples, only transitions between certain states can occur, that is,  $a_{ij} > 0$  occurs only for some  $i, j$ .

The difference between the sample and the random forest graph chain is that for each state, only one external value can be used for the observation, and the obtained observation vector is related to the state of the system, and this relationship can be discrete or continuous.

But for continuous distribution observation, the probability distribution of the observation vector corresponding to its state  $j$ :

$$b_j(v_t) = P[v_t | q_t = S_j], 1 \leq j \leq N. \tag{28}$$

Generally, the probability distribution is taken as the mixed Gaussian distribution, namely

$$b_j(v_t) = \sum_{m=1}^M \omega_{j,m} N(o_t, \mu_{j,m}, \Sigma_{j,m}). \tag{29}$$

M represents the number of mixed Gaussian distributions, and  $\bar{w}$  represents the positive mixture weight, but use  $N(o_t, \mu_{j,m}, \Sigma_{j,m})$  to represent the n-dimensional Gaussian distribution.

$$\pi_i = P[q_1 = S_i], 1 \leq i \leq N. \tag{30}$$

Therefore, the samples can be summarized into three groups  $\lambda=(A, B, \pi)$ . Then the observation sequence generated by this model can be expressed as  $O = o_1 o_2 \dots o_T$ ,  $o_t$  represents the vector that can be observed at time  $t$ , and  $T$  represents the total observation length.

*2.1.1. Digital Protection and Exploitation.* The time complexity and space complexity of high-performance computing of intangible cultural heritage are mainly affected by the dimension of the feature vector of intangible cultural heritage. In order to effectively improve the protection accuracy of the constructed model, the extraction of the model space can be determined by reducing the dimension of the intangible cultural heritage. In this paper, the information gain algorithm is used to calculate the average amount of information of the constructed model:

$$\begin{aligned}
G(w) &= - \sum_{k=1}^N P(c^k) \log_2 P(c^k) \\
&+ P(w) \sum_{k=1}^N P(c^k | w) \log_2 (P(c^k | w)) \\
&+ p(\bar{w}) \sum_{k=1}^N p(c^k | \bar{w}) \log_2 p(c^k | \bar{w}).
\end{aligned} \tag{31}$$

In the formula, represents the complementary set of  $w$ .  $\bar{w}$  corresponds to the number of spaces T: the average number of heritages in the intangible cultural heritage is trained. In the preprocessing stage, the preprocessed intangible cultural heritage is represented by word2 vec as a vector, and the intangible cultural heritage vectors with high similarity are combined to record the number of heritages. During the process of space selection in this paper, the meaning information of spatial agglomeration vectors is considered.

For built HPC, the space is traversed during state transitions. The sequence of the spatial output is expressed as using  $k$ , and  $k$  is expressed as the spatial sum that meets the similarity threshold. Therefore, in obtaining the intermediate state  $s_i$  of this model, the corresponding observation value distribution in the class  $c^k$  can be expressed as

$$b_{c_k}^i(w_{c_i}) = p\{k | s_i = w_{c_i}\}. \tag{32}$$

Taking into account that the distribution of  $b_{c_k}^i$  and spatial frequency is constrained, with the increase of spatial

distance in the whole processing process, the aggregation of different intangible cultural heritage will also be lower.

$$b_{c_k}^i(w_{c_i}) = IFI DF(i) = \frac{D_c^k(i) + 1}{\sum_c k D_c^k(i) + |C|} \times \frac{N_c^k(i) + 1}{\sum_c k N_c^k(i) + |C|}. \quad (33)$$

In the formula,  $D_c^k(i)$  represents the intangible cultural heritage business items containing  $w_{c_i}$  in  $c_k$ ;  $N_c^k(i)$  represents the number of occurrences of space  $w_{c_i}$  in the  $c_k$  category, then the number of occurrences in the same category will have The role of regularization.

For the state transition matrix, due to the high-performance computing of  $c_k$ , it can be summarized as the intangible cultural heritage in this category, then it will be transferred from the first state to the second stage state until the end of the state. The matrix of state transition  $A_c^k$  can be expressed as

$$a_{ij} = \begin{cases} 1, & j = i + 1, \\ 0, & j \neq i + 1. \end{cases} \quad (34)$$

Specify the probability  $\pi = \{1, 0, \dots\}$  of the initial state  $s_0$ . High performance computing for the category  $c_k$  can be expressed as

$$\lambda_k = \{\Pi, A_{c_k}, B_{c_k}\}. \quad (35)$$

**2.1.2. Digital Protection and Development of Intangible Cultural Heritage.** Because of the randomness of the collected data and the splitting of the node characteristics, it can ensure that there is no connection between the obtained deterministic numbers, and meanwhile, it can be used for the visual aggregation method of scientific computing, so that it has the characteristics of initial parallelism. Digital protection and development algorithms can effectively provide feedback on decision parallelization, node parallelization and feature selection.

The scientific computing visualization aggregation method is used to parallelize the collected data, which can realize the method of dispersing multiple computer nodes for the sample data during the operation. Compared with the traditional single machine identification method, it can effectively save I/O operations, and can reduce the use of more network bandwidth. According to the data obtained by parallel computing, the idea of parallelization of scientific computing visualization aggregation method is used in this paper, and combining the framework characteristics of high-performance computing, a parallel implementation strategy for scientific computing visualization aggregation is designed.

**(1) Sample Sampling Statistics of Feature Cut Points.** According to the optimal feature of the decision tree that needs to be determined, the cut point is selected for the selected feature, and the distribution feature of the data is used to sample and analyse different data features, and finally summarize to the master node of the data. This operation is mainly to perform the whole process of

computation based on different node pairs, but only a smaller bandwidth is used in the final stage of integration. Therefore, for the value of the entire eigenvalue of the data, during the process of sorting the data, the use of larger bandwidth and I/O operations should be avoided as much as possible.. The method of data collection will have a direct impact on the accuracy of the data.

**(2) Layer-By-Layer Training of Data Samples.** The deep optimization of high-performance computing in a single-sample mode is transformed into a priority strategy as the essence of each layer of training. In a distributed environment, in order to realize the training of each layer, the number of different loops that need to be constructed is the same as the maximum layer of the tree. In addition, during the loop stage, it is necessary to perform parameter statistics and judge on the split points of nonleaf nodes, according to the characteristics.

It is necessary to initialize the RFM parameters, and then use the Baum-Welch algorithm to perform operations. According to the use of the scientific computing visualization aggregation algorithm, the results obtained have a great relationship with the initial parameters. The transformation matrix needs to be initialized to determine whether the transition matrix is 0 or that after the iterative operation is 0. Define the observation sequence as

$$P(O|\lambda) \geq P(O|\lambda). \quad (36)$$

The way to compute  $P(O|\lambda)$  is a forward-backward algorithm. For the sample parameter  $\lambda$  and state  $i$ , define the forward probability  $\alpha_t(i)$ :

$$\alpha_t(i) = P(o_1 o_2 \dots o_T, q_t = i | \lambda). \quad (37)$$

That is,  $\alpha_t(i)$  is the probability that the parameter  $\lambda$  produces the sequence  $(o_1 o_2 \dots o_t)$  and the state at time  $t$  is  $o_t$ .

$P(O|\lambda)$  can be calculated by the following forward algorithm.

(1) Initialization

$$\alpha_1(j) = b_j(o_1), 1 \leq j \leq N. \quad (38)$$

(2) Recursion

$$\alpha_t(i) = b_i(o_t) \left[ \sum_{j=1}^N \alpha_{t-1}(j) \alpha_{ji} \right], 2 \leq t \leq T, 1 \leq i \leq N. \quad (39)$$

(3) Termination

$$P(O|\lambda) = \sum_{j=1}^N \alpha_T(j). \quad (40)$$

Define both  $\beta_t(i)$  and  $\xi_t(i, j)$

$$\begin{aligned} \beta_t(i) &= P(o_{t+1} o_{t+2} \dots o_T | q_t = i, \lambda), \\ \xi_t(i, j) &= P(q_t = i, q_{t+1} = j | O, \lambda). \end{aligned} \quad (41)$$

From the forward-backward calculation method,  $\xi_t(i, j)$  can also be expressed as [11, 12]:

TABLE 1: Experimental environment parameters.

Development program	Version
Operating system	Centos7.3
JDK	1.8.0-144
Internet	2 GB
Hadoop	2.7.0
High-performance computing	2.12.4
Master node	4 cores, hard disk 256G, 8G memory, one node
Worker node	1 core, 250G hard disk, 2G memory, five nodes

$$\begin{aligned} \xi_t(i, j) &= \frac{P(q_t = i, q_{t+1} = j | O, \lambda)}{P(O | \lambda)} = \frac{\alpha_t(i) \alpha_{ij} b_j(o_{t+1}) \beta_{t+1}(j)}{P(O | \lambda)} \\ &= \frac{\alpha_t(i) \alpha_{ij} b_j(o_{t+1}) \beta_{t+1}(j)}{\sum_{i=1}^N \sum_{j=1}^N \alpha_t(i) \alpha_{ij} b_j(o_{t+1}) \beta_{t+1}(j)} \end{aligned} \quad (42)$$

The probability  $\gamma_t(i, m)$  that the system is in the  $m$ -th mixed component of state  $i$  at time  $t$  is:

$$\gamma_t(i, m) = \left[ \frac{\alpha_t(i) \beta_t(i)}{\sum_{i=1}^N \alpha_t(i) \beta_t(i)} \right] \left[ \frac{\mu_{j,m} N(o_t, \mu_{j,m}, \Sigma_j, m)}{\sum_{m=1}^M \omega_{j,m} N(o_t, \mu_{j,m}, \Sigma_j, m)} \right]. \quad (43)$$

First, the labeled training set is used to train samples of various classes, respectively. Let the class that requires high-performance computing be  $k = \{1, 2, \dots, K\}$ , and each class corresponds to the model parameters  $\lambda_k$ . The maximum likelihood criterion is used to maximize the posterior probability, while considering the Bayes formula, then

$$\bar{k} = \arg \max_{1 \leq k \leq K} P(\lambda_k | O) = \arg \max_{1 \leq k \leq K} \frac{P(O | \lambda_k) P(\lambda_k)}{P(O)}. \quad (44)$$

Assuming that the prior probability  $P(\lambda_k)$  of each type is the same, because  $P(O)$  and  $k$  are not related, the decision is omitted as follows.

$$\bar{k} = \arg \max_{1 \leq k \leq K} P(O | \lambda_k). \quad (45)$$

Because  $P(O | \lambda_k)$  is very small, floating-point underflow will occur during computer operations, so logarithmic values are usually taken from it. For multiple observation sequences, only each formula needs to be weighted [14].

In the current intangible cultural heritage space identification industry, in order to meet the different needs of enterprises,  $i$  the individual needs of enterprises need to be responded to in a timely manner, and the intangible cultural heritage is mainly divided into two parts: development and design. On the basis of fully analyzing the changes in market demand, personalized services can be provided to enterprises in terms of the development process and design cycle of emerging intangible cultural heritage; meanwhile, personalized design of intangible cultural heritage can be carried out according to the actual needs of enterprises. The goals of these two links are not the same as the main tasks, but there is an internal connection. In the process of intangible cultural heritage space identification, modular and standardized design is carried out as the basis for the

intangible cultural heritage space identification process. Regarding the nonsegmented pieces of clothing in the process of intangible cultural heritage space identification, it should be transformed into a provincial virtual reality model, and it should include knowledge to support the development of new intangible cultural heritage and knowledge of corporate design according to corporate needs.

High-performance computing, which is based on virtual reality, "enhances" the real world by superimposing computer-generated virtual objects, scenes or system cues on the actual scene. High-performance computing can not only introduce virtual objects into the actual environment, but also dynamically show the position and posture of virtual objects to maintain the consistency of virtual objects and actual scenes, therefore the habit of observing high-performance computing environments with more personal eyes can be cultivated. This way of interaction looks more natural. On the other hand, since HPC keeps the actual scene, the output is more realistic [15–17].

If the high-performance computing is used, the cameras can be used to capture real-world situations to obtain video streams. The tracking method is used to process each frame in the video stream, the actual coordinates and state of the camera are calculated using geometric calculation methods, and the coordinates and states of the virtual objects registered in the actual background are used to form a virtual scene. The technology of video combination is used, to combine the virtual scene and the real background video stream are combined, and the combined result is transmitted to the display in time.

### 3. Experimental Results and Analysis

Each parallel frame in this experiment is constructed in the environment of high-performance computing cluster, the storage in the experiment is carried out by Hadoop, and the calculation of the data in the experiment is presided over by the high-performance computing program. The various development programs and their version numbers used in this experiment are shown in Table 1.

In order to verify the institutional hypothesis in the previous sentence, the verification model is set as follows.

$$ITU_{s_{ijkt}} = \alpha + \beta_1 f_{incol_{ijkt}} + \beta_2 X_{ijkt} + \vartheta_i + \tau_j + \varphi_t + \varepsilon_{ijkt}. \quad (46)$$

In the formula,  $k$ ,  $i$ ,  $j$ , and  $t$ , respectively represent the enterprise, industry, region and year;  $ITU_{s_{ijkt}}$  represents the transformation result of manufacturing enterprises;

TABLE 2: Confusion matrix.

True False	Positive		Negative
	True positive (TP)	False positive (FP)	True negative False negative

TABLE 3: Experimental results.

Experiment	High-performance computing		Traditional algorithm		Correct rate (%)
	Identify correctly	Identify errors	Identify correctly	Identify errors	
1	19	0	0	2	90.02
2	1	0	13	1	92.87
3	8	0	8	1	93.74



FIGURE 2: System interface display.

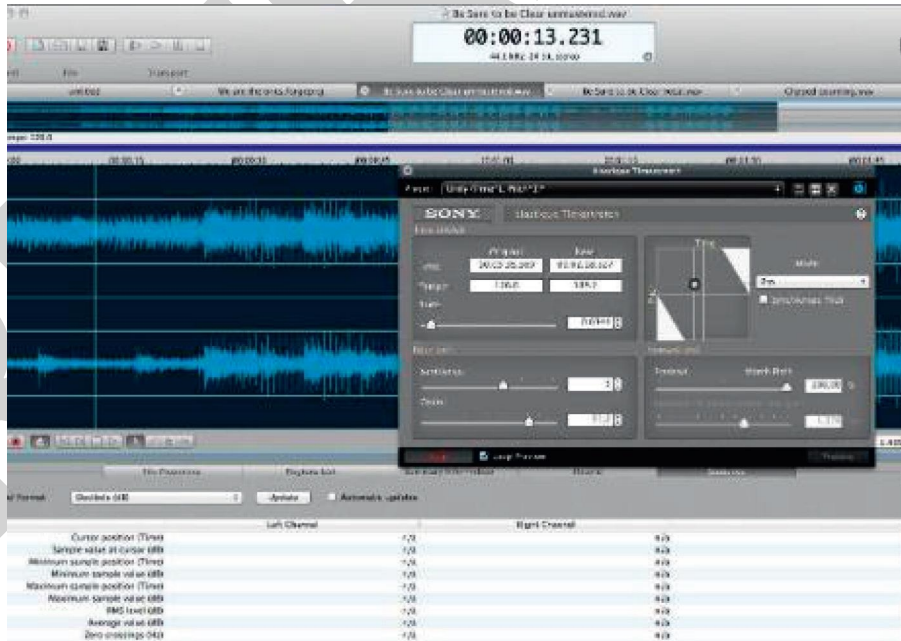


FIGURE 3: Text + audio and video augmented reality renderings.

$fincol_{ijkt}$  represents the spatial synergistic distribution between intangible cultural heritage and manufacturing;  $X_{ijkt}$  represents the control variable;  $\alpha$  represents the constant term;  $\varepsilon_{ijkt}$  is the random disturbance term. In addition,

industry fixed effects  $\vartheta_i$ , region fixed effects  $\tau_j$  and time fixed effects  $\varphi_t$  are added to the model.

The above combines the mathematical relationships between all attributes to validate the resulting attributes





FIGURE 4: Augmented reality renderings of cards before and after coloring.

obtained from the results of the classification model and evaluate the model performance against the metrics of the confusion matrix. Table 2 is the confusion matrix.

The experimental results are shown in Table 3. The experimental results show that the comparison of the accuracy of the protection of intangible cultural heritage is achieved by using the two algorithms to compare the experiments. Experiments also fully demonstrate the reliability of this method.

Figure 2 is an interface diagram of the AR module and the VR module. The module is introduced, and the interactive functions of video, voice, and intangible cultural heritage introduction are added. The origin of traditional cultural works, the production technology and development are described. Figure 3 is the effect of voice and text explanation. The VR module provides a detailed display function of VR 3D models. Users can enter the display page of VR 3D models of traditional cultural works by clicking the photos of traditional cultural works Save the model screen of interest. The AR module strengthens the relevant information of traditional cultural works through user scanning, and meanwhile paints for design on the traditional cultural works card, which enhances the interest and interaction of AR experience. This is to strengthen the display effect of the mobile terminal before and after the card is painted. In the feedback module, users can give feedback on the usage of the APP. Further enrich the functions of the APP to meet the needs of users.

In addition, the postcard design method using AR still contains traditional culture. In order to further protect the inheritance of intangible culture, the AR Postcard designed with traditional design methods according to Figure 4 can understand the traditional culture from the designed postcard. As an effective carrier of culture, the postcard can further improve the inheritance of intangible cultural inheritance, Users can use the mobile terminal to query and design ar postcards containing traditional cultural works. Ar postcards enhance the display of reality. Traditional handicrafts combine modern digital and interactive media technology to attract the younger generation to understand the production process and aesthetic characteristics of traditional handicrafts.

It is held to further publicize and popularize the protection of intangible cultural heritage. Traditional cultural works enter campus activities. Through activities in kindergartens, primary schools and other places around the city, the origin and manufacturing technology of traditional

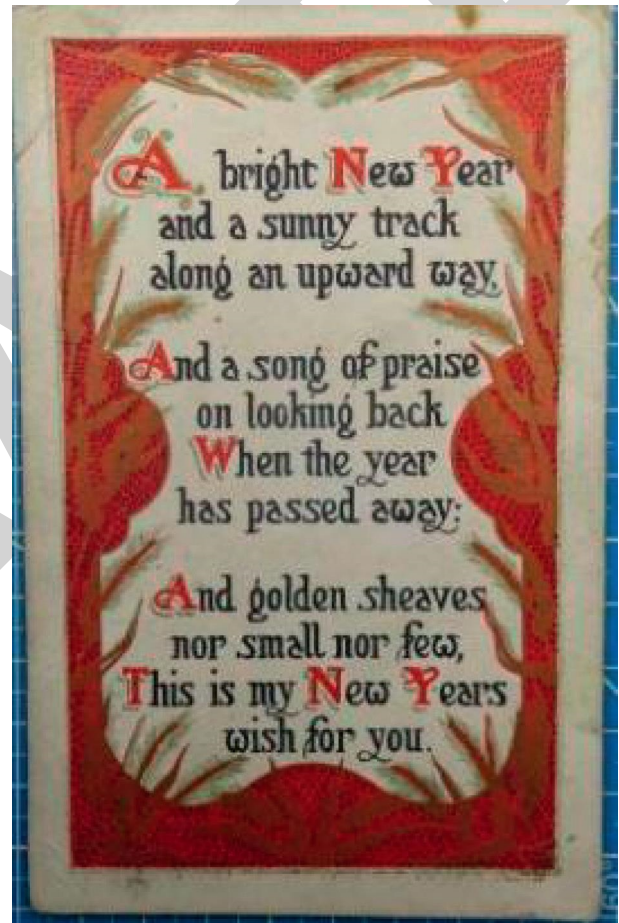


FIGURE 5: AR postcard display of traditional cultural works.

cultural works are displayed in 3D models of traditional cultural works. The mobile augmented reality APP test designed in this paper stimulates enthusiasm for learning from students and teachers to understand the intangible cultural heritage. Compared with the traditional intangible cultural protection method, this system improves the user's sense of experience and mutuality (Figure 5).

#### 4. Conclusion

In the process of the development of modern society, human's sense of inferiority towards traditional culture is not only the preservation, recording and display of cultural

## *Retraction*

# **Retracted: Cloud Computing Load Balancing Mechanism Taking into Account Load Balancing Ant Colony Optimization Algorithm**

### **Computational Intelligence and Neuroscience**

Received 17 October 2023; Accepted 17 October 2023; Published 18 October 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### **References**

- [1] J. He, "Cloud Computing Load Balancing Mechanism Taking into Account Load Balancing Ant Colony Optimization Algorithm," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3120883, 10 pages, 2022.

## Research Article

# Cloud Computing Load Balancing Mechanism Taking into Account Load Balancing Ant Colony Optimization Algorithm

Jing He 

*College of Artificial Intelligence and Big Data, Chongqing Industry Polytechnic College, Chongqing, China*

Correspondence should be addressed to Jing He; [hejing@cqipc.edu.cn](mailto:hejing@cqipc.edu.cn)

Received 19 January 2022; Revised 8 February 2022; Accepted 9 February 2022; Published 12 April 2022

Academic Editor: Daqing Gong

Copyright © 2022 Jing He. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The networking scale and traffic have exploded. At the same time, the rapid development of virtualization and cloud computing technologies not only poses a considerable challenge to the endurance of the network, but also causes more and more problems to the traditional network architecture with IP as the core. Cloud computing is a supercomputing model based on the Internet. With the rapid growth of network access and data traffic, the processing power and computing intensity will also increase, and a single server cannot afford the increase in business. In order to reduce network pressure and improve computing efficiency, load balancing for network computing is particularly important. This paper uses ant colony algorithm to design cloud computing load balance. The ant colony algorithm runs in the controller. According to the real-time network load situation provided by the controller, it calculates the link with the smallest load and provides a dynamic data stream forwarding strategy. The result of the experiments shows that the load-balanced ACO optimized technique can significantly provide an improved computational response. In the ACO algorithm, the average response time is about 30% lower than that in other algorithms. This shows that the use of the ant colony algorithm achieves a good optimization effect.

## 1. Introduction

Cloud computing is a relatively new technology based on software services such as electricity and computing. It does its job by providing service and feedback. Service provision and use conditions are important features that distinguish it from the previous complex. Low database efficiency will affect system performance, reduce system response time, and even more seriously cause node downtime. For the scheduling of cloud resources, the diversity of cloud components and the diversity of user needs make some nodes overloaded, affecting the use of the entire system and resources. Cloud computing can realize the sharing of resources, can process massive amounts of data, and can provide an unlimited amount of storage space. Therefore, cloud computing is applied to military, school, medicine, scientific research, and other fields. However, when cloud computing faces such a large number of users, it has to deal with a lot of data and tasks. When a large number of requests are processed at the same time, this will affect the

performance of the entire system. Therefore, how to achieve source access load balancing has become one of the most important issues in the implementation of cloud computing.

Cloud computing is already a significant area of research as a computer scientist. Load balancing technology is considered to be one of the most vital research directions in cloud computing and is a popular research direction that is constantly evolving. In cloud computing, the evolution must rely on the development of load balancing technology. So far, there is no ideal research result for downloading load balancing technology used in cloud environment, so it can be studied from different aspects. Due to its positive feedback mechanism and robustness, ant colony algorithm has been extensively studied and improved since it was proposed, especially for combinatorial optimization problems in application. Ant colony algorithm shows great advantages, so this article will use ant colony algorithm as the basis to apply it to the resource allocation problem in cloud computing.

A large amount of research has been done by domestic and international experts on the application of ant colony algorithms. Zhao et al. proposed a fault classification technique for neutral and invalid grounding systems based on binary ant colony algorithm (BACA) with fuzzy neural network (FNN). Test samples and field experimental findings indicate the high accuracy and good adaptability of the proposed fault classification method [1]. Vu et al. proposed the development of a quantum ant colony algorithm (QACA) to improve silent human transmission from dangerous areas to safe areas. Compared to the ACO (ant colony optimization) method, QACA has the ability to find better solutions faster using fewer individuals due to quantum representation and pheromone information and has a stronger potential. Experimental results show that the proposed method has more effective performance [2]. Pham et al. proposed a new algorithm that combines scene filtering procedures, precise algorithms, and parallel ant colony algorithms. Experiments show that building multiple initial solutions with reduced scene sets for the ant colony algorithm can significantly reduce the total time required to find the best solution. The proposed algorithm can be used for the automated construction of reliable control systems [3]. Jiang et al. used the ACO algorithm, which uses anaerobic digestion (AnD) for combination to increase biogas production to increase the release of organic waste from various waste sources in real time. An enrichment plant has been developed based on a virtual case study of organic waste affecting urban wastewater treatment plants and agri-food operations. Studies have shown the importance of the ACO method in supporting decision-making processes that help improve organic waste management and SWS control [4]. As for load balancing methods, there are fewer related studies. Hajiesmaili et al. proposed a traffic load balancing framework that seeks to strike a balance between network applications (such as total delivery delays) and green energy consumption. The statistical results show that the proposed traffic load balancing process can achieve adjustable compensations between the use of grid power and the total traffic delivery delay and saves a lot of network-connected energy at a certain cost [5]. Sarma et al. reviewed the program documentation proposed by current load balancing technologies. In addition, the advantages and disadvantages associated with many load balancing algorithms are also discussed, and the main challenges of these algorithms are identified, so that more efficient load balancing systems can be developed over time [6]. Akbari et al. proposed another algorithm with less complexity for the load balancing problem. The algorithm provides worst-case performance guarantees and describes a simple way to approximate user associations for a given input by a bias factor. Next, there is a small-scale algorithm for the overall optimization problem, with an algorithm based on the following approximate method, which has been used for nonextension optimization problems [7]. Zhao J proposed a new heuristic method for the physical host selection problem of the deployment request task, which is called Bayesian and clustering-based load balancing. As demonstrated by the results of the simulation, the proposed approach significantly reduces the

failures of the mission deployment incidents, improves the amount of throughput, and optimizes the outside-in service performance of the hybrid service in the cloud data center compared to existing work [8]. These studies provide a great reference for this article, but there are few studies on the application of ant colony optimization algorithms to cloud computing loads.

The article introduces the basic knowledge of cloud computing and the problem of resource allocation in the cloud computing environment. The objective function is proposed, and the improved genetic ant colony algorithm proposed is applied to the resource allocation problem. The operation steps and process of the algorithm are listed in detail. We integrate file size, file access speed, CPU performance, memory size, bandwidth, and other factors and perform data migration based on comprehensive load. The article takes the functional characteristics of the cloud computing platform as the background and, on the basis of the preliminary development of the current cloud computing platform resource management technology, further researches the elastic load balancing technology under the cloud computing platform.

## 2. Cloud Computing Load Balancing Method

*2.1. Cloud Computing.* Cloud computing is a super-computing model based on the Internet. It realizes the sharing of resources, and the resources and software in the cluster are allocated to users on demand. The goal of cloud computing is based on network technology to bring together traditional scattered computers to form a cluster. This combined cluster strengthens the processing capacity of the entire system.

The basic idea of cloud computing is that, in a distributed computing environment, each user can freely find the services he/she needs in this system quickly [9]. How to improve the average service response rate of the system has become a crucial issue in cloud computing [10]. The results of the cloud computing system are shown in Figure 1.

Cloud computing can be roughly divided into three categories, called the SPI model: software as a function, platform as a function, and infrastructure as a function. The infrastructure as a service provides the basis for platform services [11]. The platform as a service uses the infrastructure as a service to provide software as a service [12]. The platform as a service is similar to the web hosting service in the web hosting market. The corresponding relationship is shown in Figure 2.

Cloud computing has the following three characteristics:

- (1) Cloud computing is based on virtualization technology to realize rapid deployment of resources and access to services.
- (2) Cloud computing is provided to users in the form of services that are oriented to massive amounts of information and provided through the Internet.
- (3) The resources of cloud computing can be dynamically expanded and configured according to the needs of users and paid according to the actual use of



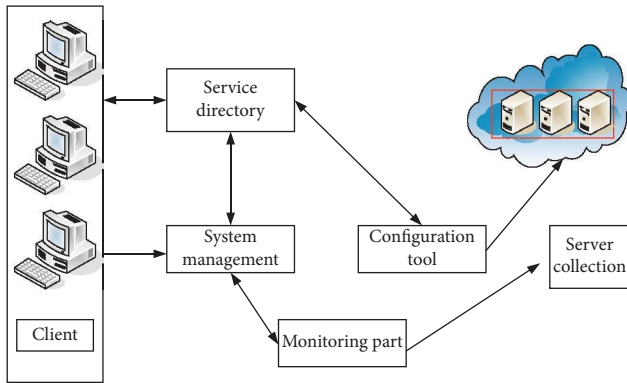


FIGURE 1: Cloud computing system.

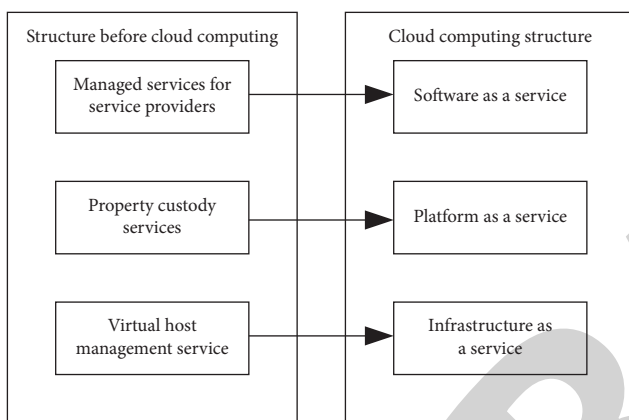


FIGURE 2: Cloud computing hierarchy.

users. Users do not need to manage them themselves, which can reduce the burden of users' processing and dependence on IT expertise.

Cloud storage uses grid technology, aggregation technology, or file sharing technology to fully organize and coordinate multiple clouds of different storage devices on each device in the cloud-covered data center [13]. Cloud storage is a cloud computing storage system. User data is stored in the cloud. For users, local storage systems and storage devices are not required [14]. More specifically, cloud storage is a service in which data is provided by a collection of multiple storage devices and servers, and users purchase services on demand.

Cloud storage deals with massive amounts of data, and it is necessary to complete the storage and management of massive amounts of data. The storage layer is composed of storage devices and storage device management systems. The storage application management system mainly completes the maintenance and upgrading of storage facilities and the monitoring of storage equipment. The structure of the cloud storage system is shown in Figure 3.

Cloud data centers usually include thousands of servers; these server nodes are heterogeneous and unevenly distributed. The storage and scheduling of resources are coordinated by many servers, and the user's access to the server is random, diversified, and complicated. The existence of

these factors can easily cause some server nodes to be idle. While other server nodes are in a state of overload saturation; that is, the load of server nodes is not balanced. The load of server nodes includes not only the load of data storage, but also the load of resource scheduling. Obviously, the existence of this imbalance will cause a waste of cloud resources, increase the burden on the network, and affect the use of users. Therefore, it is essential that the issue of load balance in relation to cloud computing is addressed. A suitable load balancing strategy allows for optimal configuration of services and increased working power.

**2.2. Load Balancing Algorithm.** Load balancing is realized by monitoring the usage of resources on each physical node and then analyzing the collected resources; by analyzing the algorithm of load balancing, the load on the node is dynamically migrated to the node with a small load; and finally the purpose of load balancing is achieved. Any computer or server source can be optimized, the response time can be as short as possible, and the processor performance can be as high as possible. Under normal circumstances, the user partition is finally obtained by analyzing the source of each controller (such as CPU usage, memory space, and bandwidth). Normally, the architecture of load balancing is shown in Figure 4.

A load balancing system is mainly composed of three parts: the cloud platform of the IaaS layer realizes the collection of resources, the realization of the load balancing strategy in the PaaS layer, and the web customer service terminal of the SaaS layer makes requests. At the IaaS layer, the underlying virtualization is mainly realized through the OpenStack cloud platform, and the monitoring function is mainly used to monitor and collect the resources of each node through Ganglia.

The PaaS layer is the core module to achieve load balancing, which is achieved through load balancing based on the dynamic load balancing algorithm mentioned in Section 3. In this layer, the dynamic scheduling of virtual machines on the nodes is realized through WSO2 software. The PaaS layer is mainly used to implement load balancing strategies and scheduling. The core processing module is the strategy controller and the queue manager. The strategy selection is mainly to analyze the load data collected by the server. By analyzing which physical machines and virtual machine resources are overloaded, underloaded, etc. and saving the data in the database, the corresponding data needs to be fetched from this database during scheduling. However, the queue manager mainly queues data.

Common load balancing algorithms include round-robin balancing algorithms, weighted round-robin symmetry, temporary balancing algorithms, and mechanical balancing algorithms.

The rotation balance algorithm used is to determine the rotation frequency before sending the request and then divide the request, rotation, etc. on the respective servers according to this process.

At the weighted round-robin balancing algorithm, the performance of each server is not considered in the round-

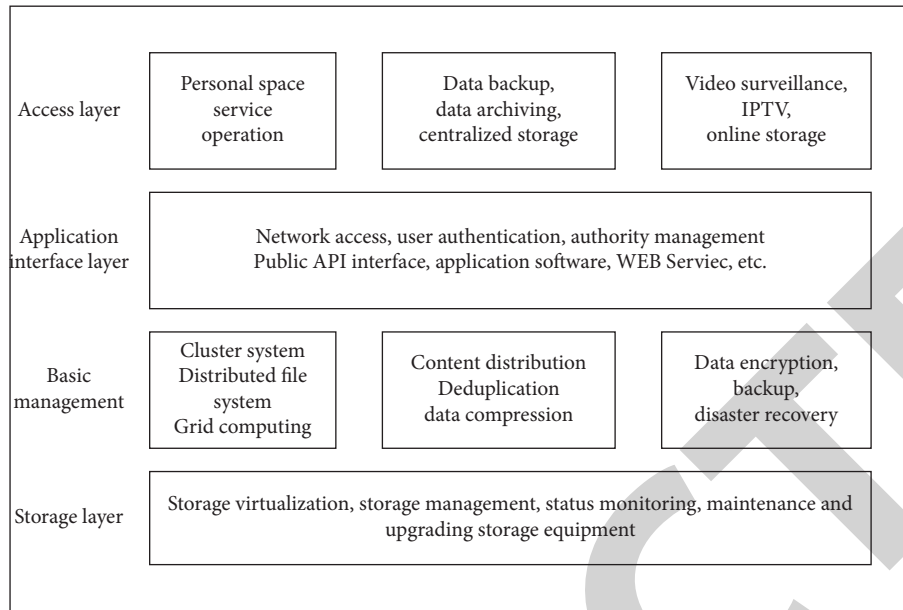


FIGURE 3: Cloud storage system structure.

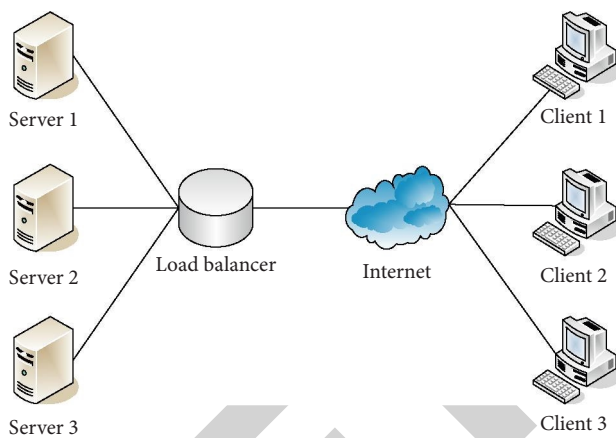


FIGURE 4: Load balancing architecture.

robin balancing algorithm. The performance of each server is inconsistent, and greater configuration and processing capacity mean greater weight; then, the load allocated to the server will be more, and the higher the load, the more efficient the operation.

The concept of load balancing is first introduced at each server node in the static payload trading algorithm, and then the corresponding function is allocated to each server, rather than the estimated resource power allocation. In the process of adjusting the concept and load balancing, the load status of each server is taken, but a value is assigned to the load, and this value will not change with the change of the load.

Consequently, the dynamic burden balancing method is a very good solution. The animated workload trading algorithm overcomes the weaknesses of static payload trading. The algorithm balances by monitoring the source status of each server and analyzing the load capacity adjustment of each node.

Through the above analysis and comparison of several load balancing algorithms, this article chooses a dynamic load balancing algorithm to achieve resource load balancing. The collection of PM (physical machine) and VM (virtual machine) data of each node in the cluster is the basis of load balancing, because algorithm balancing can be completed through the loading function of each node, and finally load balancing is determined. Because load balancing is implemented dynamically, real-time collection is required when collecting node resource information. Through analysis of the collected data, the causes of uneven load of virtual machine resources can be efficiently analyzed and processed. While collecting nodes, it is necessary to monitor the resource status of each node in real time, and the related scheduling work can be realized at the end through monitoring.

### 3. Ant Colony Algorithm

For the network, load balancing technology is the basis for the stable operation of the network and the saving of operating resources. The article will use ant colony algorithm to achieve load balancing of network links. At the same time, through the experimental analysis and selection of some parameters in the ant colony algorithm, we finally obtain a better load balancing effect.

The principle of the ant colony algorithm is that when insects carry food, they will leave a recognizable pheromone on the street, and other ant colonies can recognize this pheromone and know the concentration of the pheromone. Then, a pathway with a high pheromone concentration will be more likely to be selected, other ant colonies will pay attention to this pathway accordingly, and the pheromone concentration will increase accordingly, thus forming a positive feedback system. Therefore, the pheromone on the street will also change over time, and the pheromone concentration will decrease

accordingly, which will affect the choices of other insects. Then, in the area where the insects search for food and bring the food back to the nest, because a large number of ants continue to carry the food, it is likely that the shortest route from the nest to the foraging location will appear, and the food will eventually be collected successfully [15].

The specific expression of the ACO algorithm is shown in the following equation:

$$P_{ij}^K(t) = \begin{cases} \tau_{ij}^\alpha(t) * \kappa_{ij}^\beta(t) \\ s \in \text{allowedk} \tau_{is}^\alpha(t) * \kappa_{is}^\beta(t), s \in \text{allowedk} \\ 0, s \in \text{allowedk} \end{cases} \quad (1)$$

Among them,  $\kappa_{ij}^\beta$  is the heuristic factor of the distance perceived by the ants between  $i$  and  $j$ , and the size is related to the distance between the two cities.

$$\kappa_{ij}(t) = \frac{1}{d_{ij}}. \quad (2)$$

It can be seen from this formula that if the distance between  $(i, j)$  is smaller and the heuristic factor is greater, then the transition probability is greater. The calculation formula of  $d_{ij}$  is as follows:

$$d_{ij} = (x_i - x_j)^2 + (y_i - y_j)^2. \quad (3)$$

At all  $\tau_{ij}$  in the ACO algorithm, there are pheromone updates on the path. With the following equation, the pheromones are updated:

$$\tau_{ij}(t+1) = (1 - \gamma) * \tau_{ij}(t) + \gamma * \Delta\tau_{ij}(t, t+1). \quad (4)$$

$\gamma$  represents the pheromone volatilization coefficient between  $(0-1)$ . The specific formula of is as follows:

$$\Delta\tau_{ij}(t, t+1) = \sum_{k=1}^m \Delta\tau_{ij}(t, t+1). \quad (5)$$

Accordingly, in order to prevent the ants from traversing town  $i$  again after traversing town  $i$ , a contraposition table would be added to the ant colony algorithm to record the towns that have been traversed by the ant in a time period. After the round of traversal is complete, it will be cleared and a new round of traversal will be performed. The calculation model formula is as follows:

$$\Delta\tau_{ij}(t, t+T) = \begin{cases} K, & k \in (i, j), \\ Ln, & \text{other.} \\ 0, & \end{cases} \quad (6)$$

In the initial stage of path search, ants can choose more paths to ensure the diversity of solutions and expand the search space of solutions. Here, the  $k$ -th ant can transition from state  $i$  to state  $j$  with the following probability:

$$j = \begin{cases} \arg \max\{\tau(\gamma)^\rho\}, \\ J. \end{cases} \quad (7)$$

The article studies the mathematical principles, theories, and related applications of the ant colony algorithm. In particular, it proposes an improved ant colony algorithm for the shortcomings of the algorithm that is prone to stagnation, and finally verifies its effectiveness with specific TSPs.

Through simulation experiments, the influence of parameter range on the optimal solution is studied. The main parameters are as follows:  $\alpha$  is the relative importance of pheromone (pheromone heuristic factor);  $\beta$  is the relative importance of heuristic factor (expected heuristic factor);  $\rho$  is the pheromone volatilization coefficient ( $(1 - \rho)$  represents the persistence coefficient of pheromone);  $m$  is the number of ant colonies.

In the cloud computing environment, the problem of resource allocation is still a combination optimization problem. In the cloud computing environment, there are many factors affecting resource allocation, which correspond to the length of the intercity path in the traveling salesman problem. From the above application experiment of genetic ant colony algorithm (GACA) which improves ant colony algorithm through genetic algorithm, it is known that the fusion of intelligent algorithms is a more efficient and correct way to solve the problem. However, the merged algorithm can still have premature convergence in the later stage. Therefore, we need to maintain the diversity of the population in the iterative process and expand the search solution space to a certain extent, so as to prevent the occurrence of premature convergence to a certain extent. The definition of matching factor is introduced to express the matching degree of each task and each resource node.

$$\text{Match}_{ij} = \frac{1}{\sqrt{\sum_{n=1}^4 (T_{xi} - V_{yj})^2}}. \quad (8)$$

Supposing  $X$  is a mapping sequence of tasks and resources, the degree of load balance is defined as follows:

$$\text{Load}(X) = \sqrt{\sum_{i=1}^n \left(1 - C_1 \frac{Y_i}{Lv_i}\right)^2}. \quad (9)$$

Among them,  $Y_i$  indicates how many tasks are allocated to resource node  $i$ .

Approximate nondeterministic tree search (ANTS) is an ACO algorithm that uses mathematical programming ideas. The ANTS algorithm uses a new action selection rule and pheromone update rule, as shown in the following formula:

$$P_{ij}^k = \frac{\gamma\tau_{ij} + (1 - \gamma)\eta_{ij}}{\sum_{i \in N_i^k} \gamma\tau_{ij} + (1 + \gamma)\eta_{ij}},$$

$$\tau_{ij} \leftarrow \tau_{ij} + \sum_{k=1}^m \Delta\tau_{ij}^k, \quad (10)$$

$$\Delta\tau_{ij}^k = \begin{cases} \theta \left(1 - \frac{C^k - LB}{L_{\text{avg}} - LB}\right) \\ 0 \end{cases}$$



Among them,  $\theta$  is a parameter, and LB is an optimal solution calculated at the beginning of the algorithm.

For any given data, there is a uniquely determined membership function value  $\mu_A(x) \in [0, 1]$  corresponding to it, which can be expressed as follows:

$$\mu_{\tilde{A}}(x): X \longrightarrow [0, 1]. \quad (11)$$

Usually, the expression is as follows:

$$\tilde{A} = \frac{\mu_A^-(x_1)}{x_1} + \frac{\mu_A^-(x_2)}{x_2} + \dots + \frac{\mu_A^-(x_n)}{x_n}. \quad (12)$$

This kind of representation can be represented by sequential even representation:

$$\tilde{A} = \{(x_1, \mu_A^-(x_1)), (x_2, \mu_A^-(x_2)), \dots, (x_n, \mu_A^-(x_n))\}, \quad (13)$$

or using vector notation:

$$\tilde{A} = \{\mu_A^-(x_1), \mu_A^-, \dots, \mu_A^-(x_n)\}. \quad (14)$$

In the cloud computing system  $t$ , there are  $N$  tasks that need to be allocated to  $M$  nodes for processing, and the current load status of each server is different. The load balancing strategy is to find the most reasonable task allocation plan, so that the average processing time of tasks is the shortest, and the service quality of the cloud computing system is improved. The load deviation rate is used to measure the system load distribution after assigning tasks, and the load deviation rate is calculated as follows:

$$E(x_1, x_2, \dots, x_n) = \frac{1}{n} \sqrt{\sum_{i=1}^M e_i^2}. \quad (15)$$

Among them,  $M$  represents the predicted value of the resource load required on the  $i$ -th server under this task allocation plan, and the calculation formula is as follows:

$$e_i = f_i(x_i, L_i) - F(x_1, x_2, \dots, x_n). \quad (16)$$

Among them,  $f_i(x_i, L_i)$  represents the predicted value of resource load required on the  $i$ -th server under this task allocation plan, and the calculation formula is as follows:

$$f_i(x_i, L_i) = (x_i + 1) * L_i. \quad (17)$$

In the current server running state, the calculation formula for the impact of allocating a unit of computing resource on the load is as follows:

$$F(x_1, x_2, \dots, x_n) = \frac{1}{M} \sum_{i=1}^M f_i(x_i, L_i). \quad (18)$$

According to the above calculation formula, the load deviation rate takes a value in the range of 0 to 1. The smaller the value, the more balanced the load distribution of the

cloud computing system, and the better the performance of the current system [16].

## 4. Experiments and Results of Cloud Computing Load Balancing

**4.1. Experiment Preparation.** The cloud computing simulation platform is selected to conduct experimental analysis and evaluation of the cloud computing load balancing process based on virtual machine relocation. The main performance indicators are the load balance of CPU, memory, and other sources and the average operating time. By monitoring the load situation in the network in real time, the ant colony algorithm can give the forwarding link with the smallest load, so as to realize the relative balance of dynamic link load. This can optimize the configuration of limited network resources and effectively solve the problems of low network bandwidth utilization and low throughput. In order to simplify the processing and reflect the reliability of the experimental results of different strategies, in this simulation experiment, the cloud computing data center is configured with the same type of servers, virtual machines with the same parameters, and the same network environment and performs a set of the same tasks. The specific conditions are shown in Table 1.

It is measured by the load standard deviation for the CPU and memory of the system. Their corresponding utilization standard difference will be calculated by detecting the CPU or memory utilization at a given time for each host.

The level of memory load balancing is compared between layered management-based virtual machine migration strategy (MLVM) and agent-based virtual machine migration strategy (PAVM). After that, the entire CPU utilization of the system is read at different time points, and the standard deviation of the system CPU load balance is calculated. The experimental results are shown in Table 2.

Through experiments, the task response time for different task requests is counted, and the average task response time is calculated. The average task response time of the two strategies is shown in Table 3.

**4.2. Algorithm Optimal Solution.** We perform statistics on the optimal solutions for different  $\alpha$  values in the ant colony algorithm, and the results are shown in Figure 5.

The heuristic factor  $\alpha$  characterizes the relative importance of the pheromone accumulated by the ants in the process of constructing the solution for the path selection in the ant search. The larger the value of  $\alpha$ , the more likely the ants will pick the previous paths they traveled and the searching paths will be less stochastic. When  $\alpha$  takes different values, the corresponding optimal iterative optimal solution and number of iterations are shown in Figure 6. When  $\alpha = 1$ , the iterative optimal solution gets the minimum value, and then as  $\alpha$  increases, the iterative optimal solution value also increases; the number of

TABLE 1: Comparative experimental parameters.

Parametric variable	Parameter value
Host quantity	40
Host core number	2
Host memory	4 GB
Core computing speed	2 GHz
Initial number of virtual machines	250
Number of tasks	300
Total CPU demand	2500 GHz
Total memory requirements	4 GB

TABLE 2: System CPU load standard deviation at different time points.

Time/s	MLVM CPU load standard deviation	PAVM CPU load standard deviation
100	0.23	0.24
200	0.20	2.25
300	0.23	0.24
400	0.21	0.23
500	0.18	0.21
600	0.16	0.21
700	0.10	0.15
800	0.09	0.14

TABLE 3: Task response time.

Number of tasks	MLVM average response time	PAVM average response time
20	5.97	5.99
50	6.01	6.02
100	6.23	6.21
150	6.51	6.65
200	7.34	7.82
250	7.64	7.92
300	8.01	9.25
400	8.54	9.73

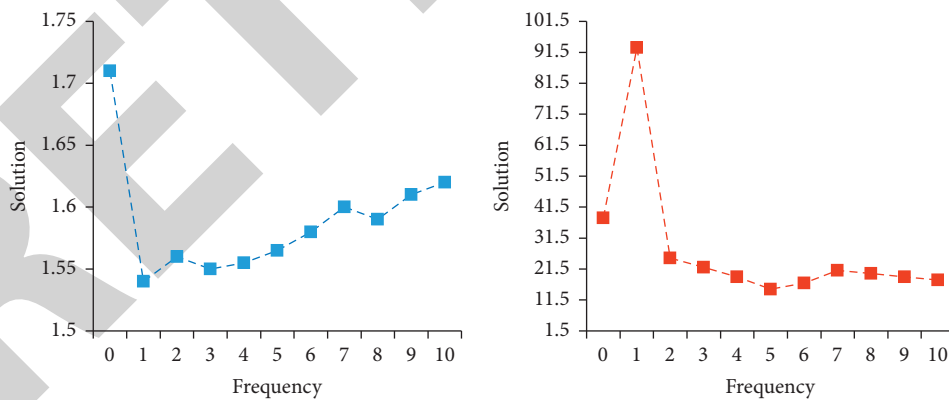


FIGURE 5: The value of the heuristic factor solution.

iterations also decreases as  $\alpha$  increases. This reflects the overall impact of  $\alpha$  on the algorithm, that is to say, the larger the value of  $\alpha$ , With the greater likelihood of selecting previously traveled paths, the more likely the ants are to choose previously walked paths; there will be less randomness in the search paths, they will converge faster, and they will be easily restricted to limiting themselves to locally optimal solutions.

We carry out statistical analysis on the load quantity of each resource of MLVM and PAVM, and the results are shown in Figure 7.

It can be seen that PAVM accepts the criteria to increase the possibility of finding a better allocation method, which is subject to certain restrictions on the execution time of the algorithm, but there is a certain improvement in the scheduling time. We compare the task time spans of several

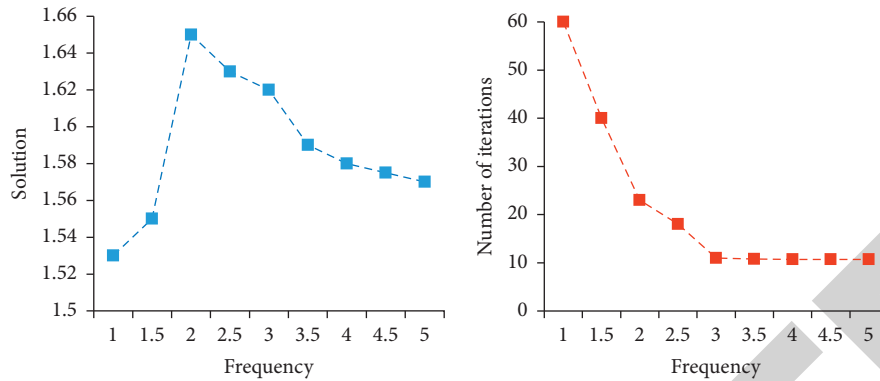


FIGURE 6: The relationship between the optimal solution and the number of iterations.



FIGURE 7: The amount of load for each resource.

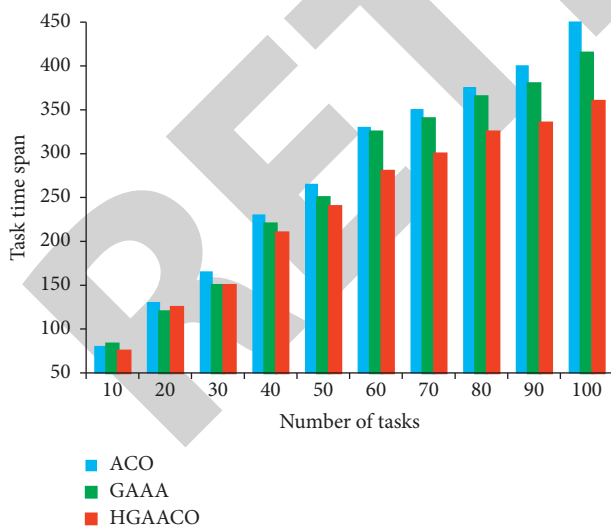


FIGURE 8: Task time span.

most commonly used algorithms, such as ant colony algorithm, genetic ant colony algorithm, and improved algorithm, and the results are shown in Figure 8.

In a load balancing system, validity is an important criterion for judging the quality of its algorithm, so it is

particularly important to verify this criterion. Due to the instability of the experimental results, the method of seeking a tie for each set of experiments is used to obtain the final result. We judged the effectiveness of different methods by the response time when reading files of different sizes. The experimental results are shown in Figure 9.

It can be seen that the response time for reading each file varies, and this time is not completely related to the file size, which means that not all large files have a long response time, and not all small files have a short response time. However, in the ACO algorithm, the average response time is about 30% lower than that in other algorithms. This shows that the ACO algorithm can play a role.

## 5. Discussion

In order to avoid data loss and improve system reliability, copy technology is usually used in cloud storage systems. The copy technology is to make multiple copies of a data block, store them on multiple nodes, and use data redundancy to ensure data reliability. The use of copy technology can well solve the load balancing problem of the system. Multiple copies are distributed on different server nodes. By accessing the copy, the load of a single server node is distributed to multiple nodes, thereby reducing the load of the nodes and achieving load balancing.

The ant colony optimization algorithm has undergone more than ten years of development since it was proposed and has achieved remarkable results in algorithm theory research and application. Its application scope involves various optimization fields such as routing problems, allocation problems, scheduling problems, subset problems, machine learning problems, and emergence of ant colony algorithm bionic hardware, showing the strong vitality and broad development of ant colony optimization algorithm prospect.

With the continuous update and development of computer technology, the number of users has doubled, the architecture of traditional IP networks has become more and more complex, and the development of the network structure is relatively lagging. In this paper, the load balance is realized by using the ant colony algorithm, which is implanted in the controller, and the network load is

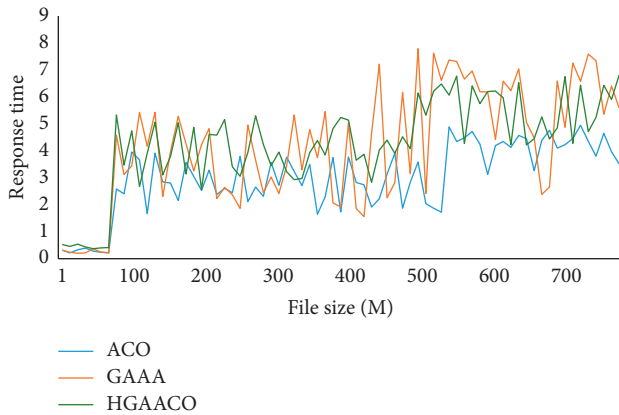


FIGURE 9: Response time for different file sizes.

monitored in real time through the traffic monitoring software. Through the execution of the ant colony algorithm in the controller, the data stream will be transmitted along the link with the least load at a certain moment, which can effectively use network resources, reduce network congestion, increase network throughput, and realize network dynamic load balancing of each link in the network. We also carried out theoretical experimental research and analysis on the heuristic factor, the pheromone volatilization coefficient, and the number of ants in the ant colony algorithm. Through the selection of parameters, the performance of the ant colony algorithm was improved, and the load balancing was also able to obtain better results. We conduct load balancing experiments on the self-defined network topology. Through the analysis of experimental data, the introduction of ant colony algorithm can achieve dynamic load balancing of each link.

## 6. Conclusion

For the purpose of this paper, we will be looking at the allocation of resources for cloud computing. According to the existing algorithm, the intelligent algorithm is fused, and the pheromone initial allocation of the ant colony algorithm is introduced to solve the blind search problem and avoid the later local optimal solution problem. The improved algorithm was applied to the specific problem of cloud computing resource allocation, and experimental calculation and analysis were performed. It was found that the time span performance and resource load degree had relatively good results, which proved the applicability of the algorithm. The improved algorithm proposed in this paper has certain performance optimization but also has certain disadvantages. The algorithm becomes slightly more complicated and has a great test on the execution speed. Moreover, the actual situation should be fully considered in the actual application process; for example, issues such as more diverse and dynamic user needs and complexity of resources require comprehensive consideration of the needs and interests of users and providers. In addition, the ant colony algorithm is often aimed at a certain type of special problem, and its versatility is poor. In future research, if we want to use the

ant colony algorithm in other fields, we need to improve its universality. In addition, the research in this paper only considers the small network with few hosts and single controller and does not consider the load balancing of the large network with multiple hosts and multiple controllers, and the researcher hopes that this part of the research can be completed through corresponding experiments in the future.

## Data Availability

The data used to support the findings of this study are available from the author upon request.

## Conflicts of Interest

The author declares no conflicts of interest.

## Acknowledgments

This research study was sponsored by Chongqing Yubei District Science and technology planning project: "Big Data Development of Service PAAS Automatic Delivery and Management Platform."

## References

- [1] J. Zhao, K. Yang, X. Wei, Y. Ding, L. Hu, and G. Xu, "A heuristic clustering-based task deployment approach for load balancing using bayes theorem in cloud environment," *IEEE Transactions on Parallel and Distributed Systems*, vol. 27, no. 2, pp. 305–316, 2016.
- [2] T. K. Vu, M. Bennis, S. Samarakoon, M. Debbah, and M. Latva-Aho, "JOINT load balancing and interference mitigation IN 5G HetNets," *IEEE Transactions on Wireless Communications*, vol. 16, no. 9, pp. 1–15, 2017.
- [3] C. Pham, N. H. Tran, C. T. Do, E.-N. Huh, and C. S. Hong, "Joint consolidation and service-aware load balancing for datacenters," *IEEE Communications Letters*, vol. 20, no. 2, pp. 292–295, 2016.
- [4] D. Jiang, P. Zhang, Z. Lv, and H. Song, "Energy-efficient multi-constraint routing algorithm with load balancing for smart city applications," *IEEE Internet of Things Journal*, vol. 3, no. 6, pp. 1437–1447, 2016.
- [5] M. H. Hajiesmaili, L. Deng, M. Chen, and Z. Li, "Incentivizing device-to-device load balancing for cellular networks: an online auction design," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 2, pp. 265–279, 2017.
- [6] A. Sarma, S. Chakraborty, and S. Nandi, "Deciding handover points based on context-aware load balancing in a WiFi-WiMAX heterogeneous network environment," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 1, pp. 348–357, 2016.
- [7] H. Akbari, P. Berenbrink, and T. Sauerwald, "A simple approach for adapting continuous load balancing processes to discrete settings," *Distributed Computing*, vol. 29, no. 2, pp. 143–161, 2016.
- [8] L. Kong, K. Ma, B. Qiao, and X. Guo, "Adaptive relay chain routing with load balancing and high energy efficiency," *IEEE Sensors Journal*, vol. 16, no. 14, pp. 5826–5836, 2016.
- [9] J. R. Aviles, M. Toril, S. Luna-Ramirez, V. Buenestado, and M. A. Regueira, "Analysis of limitations of mobility load

## Research Article

# Multidimensional Information Network Big Data Mining Algorithm Relying on Finite Element Analysis

**Haifeng Li** 

*Dalian Jiaotong University, Dalian, Liaoning 116028, China*

Correspondence should be addressed to Haifeng Li; [lihaifeng@djtu.edu.cn](mailto:lihaifeng@djtu.edu.cn)

Received 30 January 2022; Accepted 25 February 2022; Published 11 April 2022

Academic Editor: Daqing Gong

Copyright © 2022 Haifeng Li. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In recent years, with the rapid development of the Internet, online social networks have been continuously integrated with traditional interpersonal networks and research on information dissemination in social networks has gradually increased. This article studies and analyzes the multidimensional information network big data mining algorithm based on the finite element analysis method. This paper firstly introduces the finite element analysis and calculation process, a finite element data mining simulation application software management system will integrate current data, calculation, and background data into one, then analyzes the data mining clustering algorithm, and conducts an experimental exploration of the influential node mining algorithm in complex networks. The experimental results show that the LIC algorithm is better than the CC algorithm, the DC algorithm, and the BC algorithm; its overall performance is improved by 30%, and the effect is better. The LIC algorithm can effectively and quickly determine the influential nodes, which is helpful for social network analysis.

## 1. Introduction

The important purpose of the research on the mining algorithm for the number of nodes with far-reaching influence in the social Internet is to excavate the number of  $k$  important nodes with far-reaching influence in the number of nodes in the entire social Internet and finally make these  $k$  nodes in an important position. Under this transfer mechanism, as many network nodes as possible are affected. In a word, the problem of important network nodes that are discovered from social networks and have a huge social influence also constitutes a problem of maximizing social influence to some extent. Among them, mining influential nodes in the network can inhibit or accelerate the dissemination and diffusion of information in the network. It takes the mined top- $k$  nodes as a set of known seed nodes and finally maximizes the influence through a certain propagation strategy.

Due to the development of social Internet technology, the Internet has completely changed the way people communicate. Social networking on the Internet is recognized by more and more people, especially by young people, which

directly results in the rapid expansion of the social Internet. The corresponding results are roughly as follows: First, the number of user groups in the social Internet is huge, and the interpersonal network established between user groups and user groups is becoming more and more complex; the update and change speed of community Internet information content is increasing, and the relationship between users and applications is also dynamic, which makes it more difficult to predict information content. Therefore, the research and development of social Internet-oriented data mining algorithms with the characteristics of efficiency and practical value has practicality and great scientific significance for solving real problems in the community. At the same time, the research on the mining algorithm and information dissemination model of influential nodes in social network has certain scientific research value and significance for real life.

According to the research progress at home and abroad, different scholars have also carried out certain cooperative researches in finite element analysis, multidimensional information network and big data mining inside the dormitory. Elliott and Thomas defined a new finite element method for

numerical approximation of the solution of partial differential equations in the volume domain and surface partial differential equations lying on the boundary of the volume domain. The key idea is to perform a polyhedral approximation to the whole region containing the union of simplexes and to use piecewise polynomial boundary surfaces as the approximation of the surface [1]. Salonitis et al. proposes a new user-friendly optimization method for lattice component design for weight minimization, which combines finite element analysis and evolutionary calculations [2]. Li et al. analyzes the higher-order, central, finite-difference scheme of the diffusion equation in a finite interval. They analyzed stability through Gustafsson, Kreiss, and Sundström theories and eigenvalue visualization methods for semidiscrete and fully discrete schemes, followed by numerical tests to prove and validate the analytical results [3]. In order to retain important information related to mining of large data sets in multidimensional information networks, Xu presented an algorithm for constructing association rule mappings. On this basis, he proposed a multidimensional information network big data mining subalgorithm and an association rule generation algorithm, which were used for multidimensional information network big data mining and association rule generation, respectively. Finally, he demonstrated the feasibility and effectiveness of the algorithm through theoretical analysis and experimental comparison [4]. Wang proposed an indoor localization algorithm combining RSSI and nonmetric multidimensional scaling (NMDS) (RSSI-NMDS) to solve the problem that the indoor target localization algorithm based on received signal strength (RSSI) is susceptible to interference and large fluctuations in the Internet of Things environment [5]. Xu et al. takes a broader perspective on privacy issues related to data mining and studies various methods that help protect sensitive information. He identified four different types of users involved in data mining applications: data providers, data collectors, data miners, and decision makers [6]. However, these scholars did not rely on the finite element analysis method to study and analyze the multidimensional information network big data mining algorithm but only discussed its significance unilaterally.

The innovation of this article is as follows: (1) First, the finite element analysis calculation process is introduced. (2) The data mining clustering algorithm is analyzed, and the clustering analysis method refers to the analysis method of studying and processing the given object by means of computational mathematics. (3) Experimental analysis of the mining algorithm of influential nodes in complex networks is carried out.

## 2. Materials and Methods

A finite element data mining simulation uses a software management system that integrates current data, computing and background data. In the actual work of the finite element simulation analysis system, it first allows the user to use the software system to perform the preliminary data processing work of the finite element calculation with the help of the finite element data mining front-end. During this period, the

structural model must be added first, and then the automatic mesh generation program is called directly to complete the setting of the finite element mesh pattern, and finally sets important technical parameters such as material properties through communication with users [7]. Finally, a scripted document of the initial data model can be obtained as input to the finite element calculation program. In the finite element calculation process, due to the deformation caused by the finite element meshing, after the finite element data mining is completed, the meshing program must be called to obtain the calculation data. Finally, after using the program, the user can perform an in-depth study and statistical analysis of the results of the finite element calculation process. In the entire finite element calculation process, the most important is the finite element solver. Other functional modules are used for the calculation of finite element solutions, but they are also an integral part of the system. Without any module, the system cannot work properly. Figure 1 shows the service mode diagram of the finite element analysis and simulation system.

Although the finite element operation processing program is the most important part of the whole set of finite element analysis methods, the finite element operation processing program alone cannot provide an effective explanation for the actual project engineering. Therefore, on the one hand, the finite element calculation method involves many types of practical problems and a large amount of data and information. In the manual method to complete the data processing, the workload is large, the work efficiency is reduced, and errors are prone to occur; on the other hand, because the amount of information generated by the finite element operation is large and complicated, it is quite troublesome to analyze and use [8]. The finite element pre- and postprocessing software is a software technology that has flourished under this historical background.

Behavior is a series of purposeful, organized activities. There are usually two types: extrinsic behavior and intrinsic behavior. Learning behavior refers to the sum of activities that students undertake to achieve a certain learning effect under the guidance of a certain motivation. It is usually a series of activities in which students interact with their surroundings. Online learning behavior refers to the online learning that students conduct around certain learning goals, which can be set by themselves or designated by teachers. It is a collection of various explicit or implicit behaviors in learning activities. For example, uploading and downloading learning resources, learning courses on demand, publishing learning and troubleshooting, and communicating with other students [9, 10].

A large number of research results and applications in the field of data analysis and mining can be regarded as an important result of the natural evolution of technology. Strictly speaking, "data mining" usually refers to a broad concept in the industry, which should be more accurately called "mining knowledge from data" or "discovering knowledge from data"; the "data mining" in the strict sense is only a basic work stage in the entire scientific research workflow [11]. The general process of mining knowledge from data is shown in Figure 2.



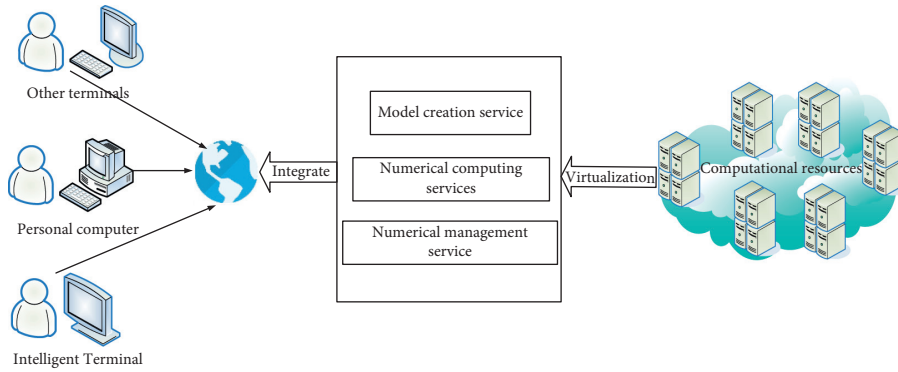


FIGURE 1: Service mode diagram of the finite element analysis simulation system.

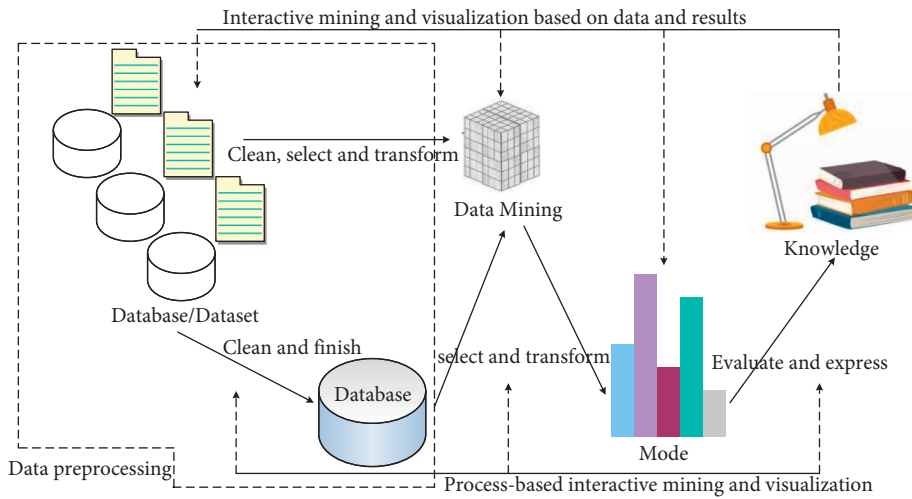


FIGURE 2: The whole process of knowledge discovery in data.

The goal of data analysis and mining is to find meaningful or potential content from the information database. The main service functions are as follows:

- (1) Concept description: concept description can also be called data synthesis. Its main purpose is to extract and compress data, and then describe the complete information, or compare it with other objects. Summary Statistics gives a general overview of the statistics.
- (2) Analysis methods and predictions: classification methods and predictions are two completely different forms of statistical analysis data, which can be used to establish models to describe important statistical categories and to predict the future development of analysis methods [12]. In practical applications, classification prediction is widely used. For example, a classification model can be established to classify mobile users and return visits to high-frequency attrition groups as soon as possible to reduce user attrition; it is also possible to classify traffic flow information by establishing a classification model to predict future traffic congestion. Cluster analysis: The basic principle of cluster

analysis is high cohesion and low mutual coupling. The final conclusion is that it has a high similarity to the same class of objects and a low similarity to other class objects [13].

- (3) Lonely point analysis: some objects in the database system may be different from the general behavior and model of data analysis, for this kind of object, we call it a lonely point. When people design data mining algorithms, they usually try to minimize the harm of lonely points, but in some special applications, the lonely point itself can also be an important judgment information. For example, in telecommunication spoofing checks, lonely spots can indicate spoofing activity [14].
- (4) Time series analysis: in time series analysis, the attribute values of data are constantly changing with time. These data are generally obtained at the same time interval but may be obtained at different time intervals. Data can be more visualized with time series plots.

Clustering refers to the process of distinguishing data signals according to specific conditions and rules. In this step, there is no prior knowledge of the relevant type, and

there is no teacher's guidance, so it is in the scope of unsupervised classification.

Cluster analysis method refers to an analysis method that uses computational mathematics to study and process a given object. Clustering is an important and human act because a person's developmental process learns to recognize things by constantly reshaping subconscious aggregation patterns. The traditional cluster analysis method strictly defines various statistical objects to be aggregated, and the boundaries between classifications are very clear, so we can call it hard classification [15]. However, in general practical use, because most objects lack strict characteristic attribute definitions, and there is an unclear intermediary between the form and the class characteristic attributes, they have the characteristics of one and the other, so they are suitable for implementing soft classification.

**2.1. Mathematical Model of Cluster Analysis.** Supposing  $A = \{a_1, a_2, \dots, a_m\}$  is the whole object to be clustered (i.e. domain of discourse), among them, each object  $a_i$  ( $1 \leq i \leq m$ ) is called a sample, and the object is described by a plurality of parameter values, which represent a feature of the object. So each object  $a_i$  corresponds to a vector  $V(a_i) = (a_{i1}, a_{i2}, \dots, a_{in})$ , where  $a_{ij}$  ( $1 \leq j \leq n$ ) is the assignment of  $a_i$  on the  $j$ th feature and  $V(a_i)$  is called the feature vector of  $a_i$  [16]. Cluster analysis is to analyze the spatial distance between the feature vectors corresponding to the samples in the universe  $A$  and their distribution and divide them into  $l$  disjoint pattern subsets according to the distance relationship between each sample and satisfy the following formula conditions:

$$\begin{aligned} A_1 \cup A_2 \cup \dots \cup A_l &= A, \\ A_i \cap A_j &= \emptyset \quad (1 \leq i, j \leq l, i \neq j). \end{aligned} \quad (1)$$

Membership function  $u_{ij}$  represents the membership of sample  $a_j$  ( $1 \leq j \leq m$ ) with respect to subset (class)  $A_i$  ( $1 \leq i \leq l$ ), as follows:

$$u_{a_i}(a_j) = u_{ij} = \begin{cases} 1 & (a_j \in A_i), \\ 0 & (a_i \in A_j). \end{cases} \quad (2)$$

As shown in the previous formula, the membership function must satisfy  $u_{ij} \in N_{sl}$ , that is, each sample can only belong to one of the classes, and each subclass is required to be a nonempty set.

$$N_{sl} = \left\{ u_{il} | u_{ij} \in \{0, 1\}, \sum_{j=1}^m u_{ij} = 1, \forall j; 0 < \sum_{i=1}^l u_{ij} < m, \forall i \right\}. \quad (3)$$

In the real world, the membership degree of a sample may not be either one or the other. From this, fuzzy theory is introduced, which becomes a fuzzy set class, and the sample set  $A$  is divided into multiple fuzzy subsets  $\tilde{A}_1, \tilde{A}_2, \dots, \tilde{A}_l$ , and the membership function  $u_{ij}$  of the sample is extended from the binary case of  $\{0, 1\}$  to the  $[0, 1]$  interval, satisfying the conditions:

$$\begin{aligned} \sup v(\tilde{A}_i) &= A, u_{ij} \in [0, 1], \\ \sum_{i=1}^l u_{ij} &= 1 (\forall j), 0 < \sum_{j=1}^m u_{ij} < n, \end{aligned} \quad (4)$$

where  $\sup$  represents the support set of the fuzzy set.

**2.2. Fuzzy Clustering Objective Function.** Mathematical model based on cluster analysis, if fuzzy concept is introduced. It uses the membership degree  $o_{il}$  of the sample point  $a_i$  and the sample prototype of the  $i$ -th class as the weight to weight the distance. At this time, the objective function of the cluster analysis is given as follows:

$$J_1(O, V) = \sum_{i=1}^b \sum_{l=1}^m o_{il} (h_{il})^2, \quad O \in R_s. \quad (5)$$

After fuzzy clustering the objective function of the hard clustering analysis method, it uses the sum of squares of membership to transform the objective function of the sum of squares of errors into the objective function of all weighted sums of squares of errors, namely,

$$J_2(O, V) = \sum_{i=1}^b \sum_{l=1}^m o_{il}^2 (h_{il})^2, \quad O \in R_s. \quad (6)$$

**2.3. Objective Function of General Fuzzy Cluster Analysis.** Among the more general expressions are

$$J_n(O, V) = \sum_{i=1}^b \sum_{l=1}^m o_{il}^n (h_{il})^2, \quad O \in R_s, \quad (7)$$

where  $h_{il}$  is a distance norm, and the distance  $h_{il}$  between the sample point  $a_i$  and the cluster center  $v_i$  of the  $i$ -th class represents the similarity between the sample point and the cluster center, which can generally be expressed as follows:

$$\begin{aligned} h_{il}^2 &= a_i - v_{iX} \\ &= (a_i - v_i)^Q X(a_i - v_i), \end{aligned} \quad (8)$$

where  $n \in [1, \infty)$  is the weighting index, also known as the smoothing parameter, which controls the fuzzy degree of fuzzy clustering. The larger the  $n$  is, the larger the blurring degree is; the smaller the  $n$  is, the smaller the blurring degree is [17]. Because  $n$  controls the degree to which membership is shared among classes, the larger  $n$ , the greater the ambiguity. After research, it is found that for different application scenarios, the optimal  $n$  can be in the range of 1 to 5, and  $n = 2$  is usually selected.

**2.4. Analysis and Solution Process Based on Objective Function Clustering Algorithm.** In order to obtain the optimal solution of the objective function of fuzzy clustering, the criterion of clustering can be taken, that is,  $\sum_{i=1}^b o_{il} = 1$  and  $\min\{J_n(O, V)\}$  are obtained under the constraints of the mechanism, namely,

$$\begin{aligned} \min\{J_n(O, V)\} &= \min\left\{\sum_{l=1}^m \sum_{i=1}^l (o_{il})^n (h_{il})^2\right\} \\ &= \sum_{l=1}^m \min\left\{\sum_{i=1}^b (o_{il})^n (h_{il})^2\right\}. \end{aligned} \quad (9)$$

Therefore, this problem can be understood as follows: under the condition of membership degree,  $\sum_{i=1}^b o_{il} = 1$ , obtain

$$\min\left\{\sum_{i=1}^b (o_{il})^n (h_{il})^2\right\}. \quad (10)$$

Using Lagrange's theorem to solve, we can get the expression of the membership degree corresponding to the  $j$ th object:

$$o_{il} = \frac{1}{\sum_{i=1}^b [h_{jl}/h_{il}]^{2/n-1}}. \quad (11)$$

The fuzzy classification matrix usually satisfies the following conditions:

- (1)  $o_{il} \in [0, 1]$ , that is, each element in the matrix is between 0 and 1 closed interval
- (2)  $\sum_{i=1}^b o_{il} = 1$ , the sum of the elements of each column is 1, that is, the sum of the membership degrees of each element to all classes is 1
- (3)  $\sum_{i=1}^b o_{il} > 0$ , this condition ensures that each class is not empty

Similarly, the cluster center of fuzzy clustering can be obtained as follows:

$$V_i = \frac{\sum_{l=1}^m (o_{il})^n a_l}{\sum_{l=1}^m (o_{il})^n}. \quad (12)$$

It can be seen that the cluster center can be obtained from the membership degree. The purpose of cluster analysis is to discover the implicit grouping information in the data. In order to represent this accuracy well, the effectiveness of clustering needs to be analyzed. When evaluating clustering results, the optimal number of clusters is usually an important indicator of effectiveness [18, 19].

In the process of optimizing the objective function, people have tried many methods such as dynamic programming, but the large storage space and long running time limit its popularization and application. The most widely used in practice is the iterative optimization algorithm, which is easy to enter the extreme point of the local area, so it is more sensitive to initialization, and one of the important input parameters is the number of clusters  $k$ . Although the exact value of  $k$  may not be known, it is usually possible to approximate the value of  $k$  by evaluating the quality of the clustering results at different  $k$ . Given an ideal cluster index, such as average radius or diameter, as long as the number of an imaginary cluster is less than or greater than the actual number of clusters, the index will tend to rise very slowly [20]. However, when trying to acquire far fewer clusters than

the true number, the index increases rapidly. A schematic diagram of the above idea is shown in Figure 3.

Among them, for the evaluation of fuzzy clustering, the separation coefficient is expressed as follows:

$$G(O, h) = \frac{1}{m} \sum_{i=1}^m \sum_{j=1}^h (o_{ij})^2, \quad (13)$$

where  $o_{ij}$  is the fuzzy membership degree of data object  $i$  belonging to cluster  $j$ , assuming all clustering results here, then the number of clusters  $k$  is given by the following formula:

$$\max\{\max G(O; h)\}, h = 2, 3, \dots, m-1. \quad (14)$$

Similarly, a variant of the separation coefficient can be obtained, the separation entropy index:

$$L(O, h) = -\frac{1}{m} \sum_{i=1}^m \sum_{j=1}^h o_{ij} \log(o_{ij}). \quad (15)$$

The corresponding number of clusters is given as follows.

If  $o_{ij}$  is close to 0 or 1, the smaller the entropy value and the better the clustering result. If  $o_{ij}$  is close to 0.5, the clustering ambiguity is high, the entropy value is large, and the corresponding clustering result is not good. These two indicators are proposed for partition-based fuzzy clustering, which is sensitive to fuzzy factors [21].

In the fuzzy k-means algorithm, due to the introduction of  $\sum_{i=1}^b o_{ih} = 1$ , in the case that the sample set is not ideal, it can be assumed that there are several isolated points far away from the clustering center of each class. Because the isolated point itself actually belongs to a certain class, under the original constraints, it will have a large degree of membership to various types or be approximately equal to each type of membership degree. This situation will affect the correctness of the final clustering results [22]. To overcome this shortcoming, we can relax the constraint and change it to

$$\sum_{i=1}^b \sum_{h=1}^m o_{ih} = \beta, \beta \in [1, m]. \quad (16)$$

That is, the sum of the membership degrees of all samples to each category is  $\beta$ , a positive integer value. At this time, on the premise of Formula (16), obtain  $\min\{J_n(O, V)\}$ , and the solution method is the same as above. At this point, the membership degree should be

$$o_{ih} = \beta \frac{1}{\sum_{i=1}^b \sum_{i=1}^m [s_{jh}/s_{ih}]^{2/n-1}}. \quad (17)$$

Obviously, this condition is completely different from the membership function in the traditional sense, and the membership degree estimated by the improved fuzzy k-means algorithm may exceed 1. If necessary, the obtained membership value can be normalized, which will not affect the final clustering result under normal circumstances.

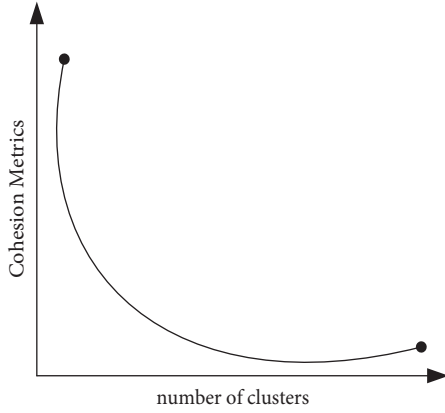


FIGURE 3: Trend diagram of the number of  $k$  versus the average diameter of clusters.

The improved method has better adaptability than the original algorithm. It can not only obtain better processing results for clustering in the case of special isolated points but also relax the constraints of membership degree. As a result, the original result is less sensitive to the predetermined number of clusters than the clustering result. When the number in the cluster analysis method determined in advance is much lower than the value of the theoretical number of clusters, the fuzzy k-means calculation using this method can obtain a more accurate cluster analysis method center result.

However, if a certain cluster center is very close to a certain sample in the iterative process, then a cluster containing only one sample may be obtained in the end. To prevent this from happening, a nonlinear process can be added to the distance operation in  $o_{ih}$ , so that the minimum calculated distance is not less than a certain threshold.

The usual distance measure in classical k-means clustering algorithm and fuzzy k-means clustering algorithm adopts Euclidean distance or Manhattan distance. The Kowski distance is a generalization of these two distances, and it is defined as follows:

$$s(i, j) = \left( |a_{i1} - a_{j1}|^t + |a_{i2} - a_{j2}|^t + \dots + |a_{iv} - a_{jv}|^t \right)^{1/t}. \quad (18)$$

When  $t = 1$ , it is the Manhattan distance, and when  $t = 2$ , it means the Euclidean distance.

In cluster analysis, often fields in a data set may maintain an established or obvious weighting relationship. For example, in the cluster analysis of mobile customers, the age, income, and other classification weights of customers are significantly larger, while the weight of noise points is significantly smaller. So, we get the weighted Euclidean distance:

$$s(i, j) = \left( u_1 |a_{i1} - a_{j1}|^2 + u_2 |a_{i2} - a_{j2}|^2 + \dots + u_v |a_{iv} - a_{jv}|^2 \right)^{1/2}. \quad (19)$$

The advantage of this method is that it can find more accurate cluster centers than the original algorithm, the disadvantage is that it is difficult to determine the size of the

weights. In order to realize the automation and intelligence of weight judgment, it can consider weighting each sample through the improved membership function. The weights of noise samples after relaxing the membership constraints are usually small, so the expected purpose is achieved. Then, the minimum value of the objective function is obtained to realize cluster analysis. Figure 4 shows the cluster analysis process.

### 3. Results and Analysis

**3.1. Influential Node Mining Algorithm for Complex Networks.** The evaluation of the influence of nodes in complex networks can make information spread faster and more widely, which is of great significance in practical applications. According to the characteristics of complex networks, this article proposes a Local Importance Centrality Algorithm (LIC), which is a new method for evaluating the influence of nodes in complex networks based on network local characteristics, such as degree and clustering coefficient. The local importance centrality algorithm can more effectively identify influential nodes.

**3.2. Experimental Data Set.** The data sets used in this experiment are collected from real networks: Netscience, E-mail, and Power.

Netscience data set publishes a topological graph of collaborations between authors of articles on topic networks, in which scientists represent nodes and collaborations represent edges. The data set contains a total of 1589 scientists, using the largest subgraph with 379 nodes in it. The basic data parameters of the data set are shown in Table 1.

E-mail data set represents the topology of e-mail forwarding relationships among Rovira University members, where Rovira University members represent nodes and forwarding relationships between members represent edges. The data parameters are shown in Table 2.

Power data set contains an undirected and unweighted topology map of the National Grid in the western United States. Nodes represent transformers, substations, and generators, and the relationship is a high-voltage transmission line. The basic data parameters of the network are shown in Table 3.

#### 3.3. Experimental Results and Analysis

**3.3.1. Relationship between the Rank Evaluated by Various Centrality Algorithms and the Average Influence Value  $\langle F(t) \rangle$ .** In this article, the SIR infectious disease model is used to simulate the propagation on three real network data sets, and the top-rank of the influence ranking obtained by the DC algorithm, the BC algorithm, the CC algorithm, and the LIC algorithm is compared experimentally.

Only one node is selected as the initial infection node for each execution of influence propagation, and other nodes are susceptible nodes. Then, the information or virus spreads according to the mechanism of the SIR infectious disease model. After  $n$  executions (each node has one and only one

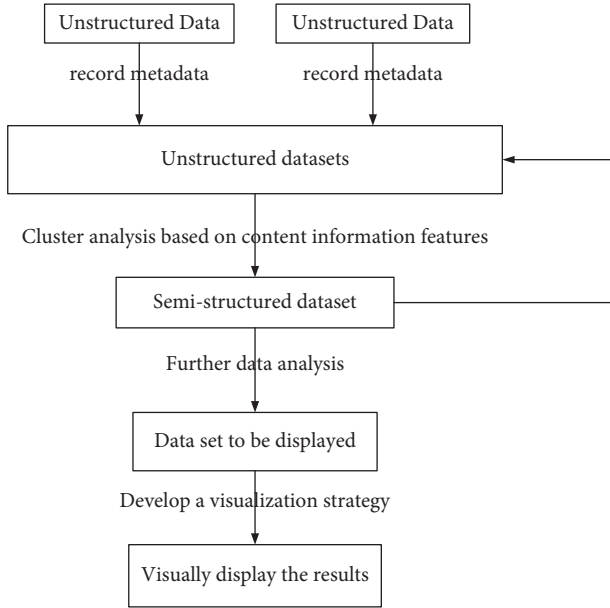


FIGURE 4: Cluster analysis.

TABLE 1: Basic network parameters of the Netscience data set.

Statistical item	Statistics	Statistical item	Statistics
Number of nodes ( $n$ )	381	Maximum degrees	36
Number of sides ( $m$ )	916	Clustering coefficient	0.39
Average degrees	4.79	Average shortest distance	5.98

TABLE 2: Basic network parameters of E-mail data set.

Statistical item	Statistics	Statistical item	Statistics
Number of nodes ( $n$ )	1098	Maximum degrees	69
Number of sides ( $m$ )	5469	Clustering coefficient	0.13
Average degrees	9.57	Average shortest distance	3.697

TABLE 3: Basic network parameters of the Power data set.

Statistical item	Statistics	Statistical item	Statistics
Number of nodes ( $n$ )	5021	Maximum degrees	20
Number of sides ( $m$ )	6603	Clustering coefficient	0.13
Average degrees	2.58	Average shortest distance	19.02

initial infection node), the correlation between the node influence value  $F(t)$  and the four centrality algorithms can be examined. In order to further clarify the performance of several centrality algorithms, calculate the average influence value  $\langle F(t) \rangle$  of the top- $k$  nodes of each algorithm, where  $\langle F(t) \rangle$  is the average of all  $F(t)$ s after  $n$  executions.

Comparisons are made on the Netscience data set. As shown in Figure 5, if the propagation performance of the centrality algorithm is good, then the top- $k$  ranking Rank and  $\langle F(t) \rangle$  should be negatively correlated, and the curve is downward sloping. That is, as the Rank increases, the nodes can influence (infect) fewer nodes.

Figure 6 shows the curves of the four algorithms are all downward sloping, and the curve of the degree centrality algorithm (DC) fluctuates greatly, which shows that the Rank calculated by the DC algorithm is quite different from the actual influence value. The curve of the DC algorithm is at the bottom of the four curves, which further shows that the top- $k$  nodes discharged by the DC algorithm have a relatively low influence. The betweenness centrality algorithm (BC) also has a large curve and performs worse than the closeness centrality algorithm (CC) and the locally important centrality algorithm (LIC). The performance of the LIC algorithm is slightly better than that of the CC algorithm, and the two curves are relatively close.

On the E-mail data set, Figure 6 shows that the experimental results of the four centrality algorithms in this data set are all good. The top- $k$  nodes sorted by the four algorithms can affect many nodes. It can be seen that the effect of the LIC algorithm is the best among the four centrality algorithms, and its curve fluctuation is relatively smooth. The effect of the CC algorithm in the figure is also very good. The  $\langle F(t) \rangle$  of the BC algorithm and the DC algorithm are very high, but the curve fluctuations of the two are relatively large and the correlation is poor.

**3.3.2. Comparison of the Influence of Top 10 Nodes Based on the Sorting of Several Centrality Algorithms.** In this round of experiments, the random sorting algorithm was added for comparison. This article compress the influence  $F(t)$  of different nodes in the top-10 node list evaluated by these five algorithms. Because the same nodes in the top-10 list are excluded, the performance of each algorithm can be compared better. Figures 7 and 8 show that on the Netscience data set, the LIC algorithm is compared with the BC algorithm, the CC algorithm, the DC algorithm, and the random algorithm. The total number of infected and immune nodes increases with time  $t$  and eventually reaches a stable value. It can be seen that although the convergence speed of the LIC algorithm is slightly lower than that of the CC algorithm and the DC algorithm, it can infect more nodes than the two algorithms. The LIC algorithm can obviously infect more nodes than the Random algorithm.

Figure 9 shows the comparison of the influence  $F(t)$  of the LIC algorithm and the DC algorithm at time  $t$  on the E-mail data set and the Power data set, respectively. The convergence speed of the LIC algorithm is significantly higher than that of the DC algorithm.

In Table 4, the average  $F(t)$  value of the top-10 nodes of the four algorithms is calculated. It can be seen that the average value of  $F(t)$  of the LIC algorithm is slightly better than that of the CC algorithm, its performance is improved by 30%, and the effect is better, and both are better than the DC algorithm and the BC algorithm.

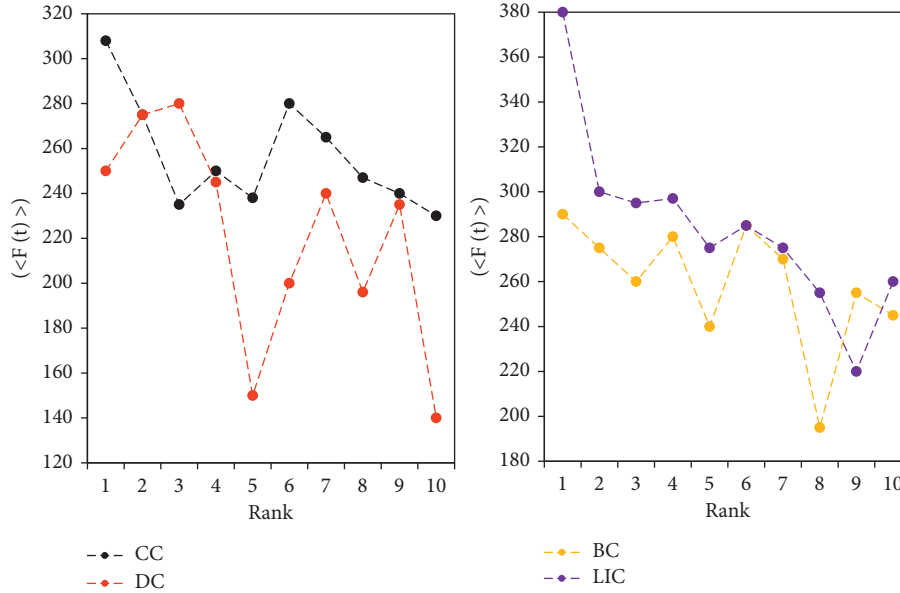


FIGURE 5: Comparison of the correlation between Rank and  $\langle F(t) \rangle$  on the Netscience data set.

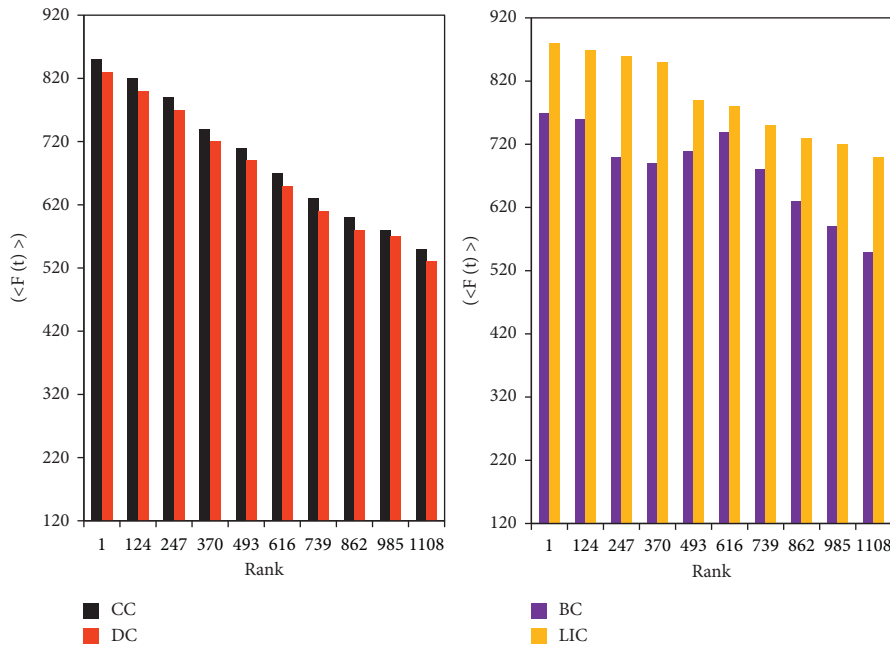


FIGURE 6: Correlation comparison between Rank and  $\langle F(t) \rangle$  on the E-mail data set.

This article proposes a new centrality algorithm based on network local importance for mining the influence of nodes in social networks. To verify the effectiveness and superiority of the algorithm, it uses the SIR infectious disease model to simulate the actual spreading influence of sorted list nodes evaluated by different centrality algorithms. The experimental results on three real social network data sets show that the influence nodes mined by the local importance centrality algorithm (LIC) are significantly better than the betweenness centrality algorithm (BC) and the degree centrality algorithm (DC). The influence effect is close to the closeness centrality algorithm (CC) and is much lower than

the closeness centrality algorithm (CC) in computational time complexity. This shows that the LIC algorithm can solve the shortcomings and deficiencies of the existing centrality algorithms and can effectively and quickly determine the influential nodes, which is helpful for social network analysis and has important practical significance.

#### 4. Discussion

The current research results show that some network nodes in social portals have great social influence, which can promote or limit the dissemination of information to



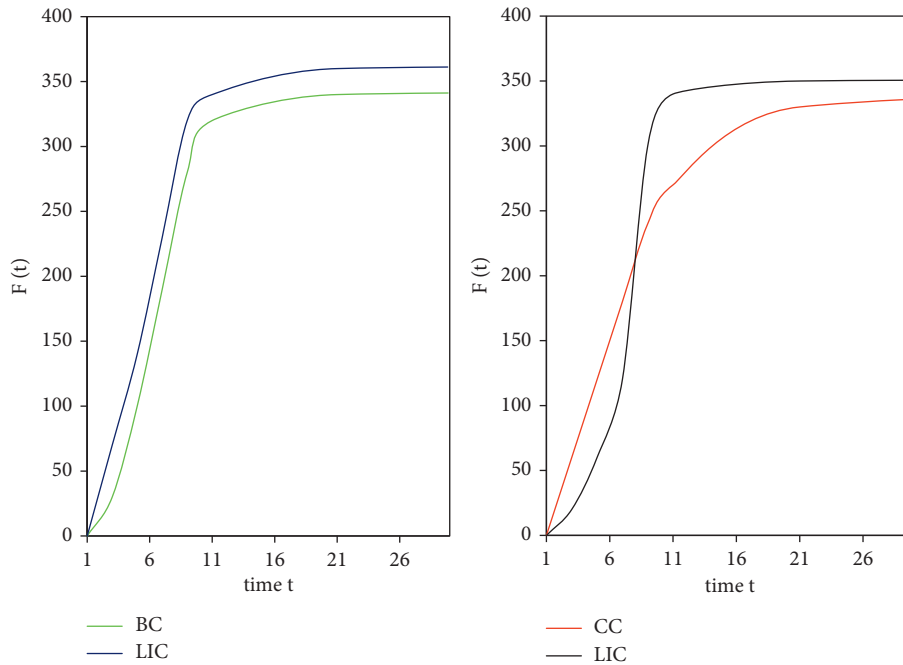


FIGURE 7: Comparison of the number of infected nodes between the LIC algorithm and other algorithms in time  $t$  (1).

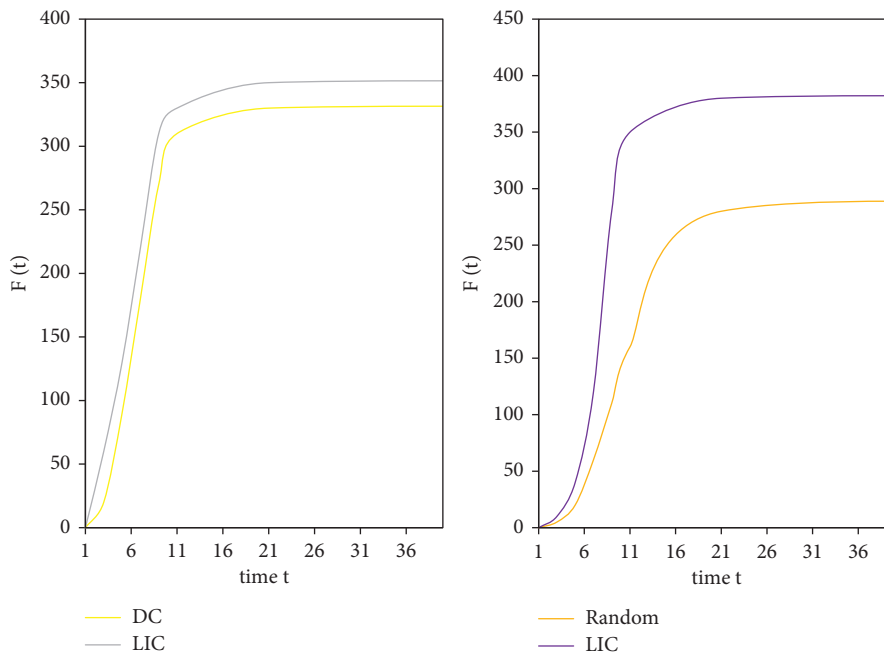


FIGURE 8: Comparison of the number of infected nodes between the LIC algorithm and other algorithms in time  $t$  (2).

a certain extent. These network nodes reduce the energy consumption of traditional Internet computing resources to a certain extent and give full play to the function of effectively optimizing traditional Internet information resources. Therefore, finding and classifying influential network nodes has become an urgent problem for many experts and scholars. The characteristic of the social network analysis method compared with the other

two methods is to analyze the network characteristics without destroying the network structure. This method is universal and suitable for most social networks. The social network analysis method mainly relies on analyzing the basic topological characteristics of the network, such as the degree of the node, the shortest path, the eigenvector, and other indicators to evaluate the influence of the node.

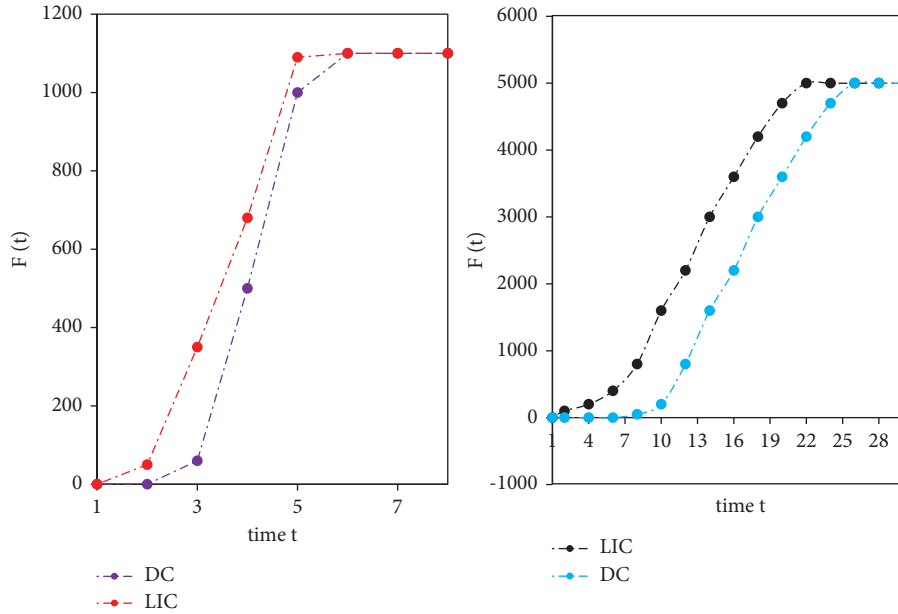


FIGURE 9: Comparison of the influence  $F(t)$  between the LIC algorithm and the DC algorithm at time  $t$ .

TABLE 4: Average  $\langle F(t) \rangle$  values of top 10 nodes for each centrality algorithm.

Data set	LIC	DC	BC	CC
Netscience	359.8	358.7	357.3	358.3
E-mail	1009.8	996.2	1005.5	1010
Power	1399.6	589.1	1389.7	1393.2

## 5. Conclusion

In this article, the influence of nodes is approximated by the local importance of nodes in the network, and a new algorithm based on the local importance centrality of the network (LIC algorithm) is proposed. The local importance of nodes is calculated and evaluated according to local information such as network moderateness and aggregation coefficient, so the computational time complexity is relatively low. By comparing it with the degree centrality algorithm (DC), the betweenness centrality algorithm (BC) and the closeness centrality algorithm (CC), it proves that it has a high influence correlation with the local important centrality algorithm, so that important influential nodes can be well mined. In this article, the research on mining key nodes in the social Internet is mainly based on the current more common way to evaluate the influence of nodes through node centrality. But in real, online, social networking sites, users' influence is not only evaluated by topological characteristics among users, but some more personalized characteristics, such as user attributes, preferences, and the like. Therefore, the next research work is to add attributes such as user personalization characteristics on the basis of complex network topology characteristics to evaluate user influence.

## Data Availability

The data used to support the findings of this study are available from the author upon request.

## Conflicts of Interest

The author declares no conflicts of interest.

## Acknowledgments

This research study was sponsored by China Association of Higher Education 2020 Special Project (2020XXHYB09) and Project of Liaoning Provincial Education Department (JDL2019027). The author thanks these projects for supporting this article.

## References

- [1] C. M. Elliott and R. Thomas, "Finite element analysis for a coupled bulk-surface partial differential equation," *IMA Journal of Numerical Analysis*, vol. 33, no. 2, pp. 377–402, 2018.
- [2] K. Salonitis, D. Chantzis, and V. Kappatos, "A hybrid finite element analysis and evolutionary computation method for the design of lightweight lattice components with optimized strut diameter," *International Journal of Advanced Manufacturing Technology*, vol. 90, no. 9-12, pp. 2689–2701, 2017.
- [3] T. Li, C.-W. Shu, and M. Zhang, "Stability analysis of the inverse lax-wendroff boundary treatment for high order central difference schemes for diffusion equations," *Journal of Scientific Computing*, vol. 70, no. 2, pp. 576–607, 2017.
- [4] D. Xu, "A large data mining for multidimensional information network relying on association rule mapping," *Boletim Tecnico/Technical Bulletin*, vol. 55, no. 17, pp. 392–400, 2017.

- [5] S. Wang, "Wireless network indoor positioning method using nonmetric multidimensional scaling and RSSI in the Internet of things environment," *Mathematical Problems in Engineering*, vol. 2020, no. 1, 7 pages, Article ID 8830891, 2020.
- [6] L. Xu, C. Jiang, J. Wang, J. Yuan, and Y. Ren, "Information security in big data: privacy and data mining," *IEEE Access*, vol. 2, no. 2, pp. 1149–1176, 2017.
- [7] E. M. Imran, R. Khan, and S. Badshah, "Finite element analysis to investigate the influence of delamination size, stacking sequence and boundary conditions on the vibration behavior of composite plate," *Iranian Journal of Materials Science & Engineering*, vol. 15, no. 1, pp. 11–21, 2018.
- [8] N. Sansone Ruiz, "Arguing against graphic ambivalence: what earth modeling reveals about visualization in scientific computing," *Information & Culture*, vol. 55, no. 3, pp. 204–225, 2020.
- [9] J. Nonaka, E. C. Inacio, K. Ono, F. Shoji, Y. Kawashima, and M. Dantas, "Data I/O management approach for the post-hoc visualization of big simulation data results," *International Journal of Modeling Simulation & Scientific Computing*, vol. 9, no. 3, pp. 1840006.1–1840006.16, 2018.
- [10] P. K. Dutta and S. Banerjee, "Monitoring of aerosol and other particulate matter in air using aerial monitored sensors and real time data monitoring and processing," *Journal of System and Management Sciences*, vol. 9, no. 2, pp. 104–113, 2019.
- [11] D. Ganellari, G. Haase, and G. Zumbusch, "A massively parallel Eikonal solver on unstructured meshes," *Computing and Visualization in Science*, vol. 19, no. 5-6, pp. 1–16, 2018.
- [12] M. Michalski, M. Rieth, A. Kempf, and J. Kruger, "CoFlaVis: a visualization system for pulverized coal flames," *Computing in Science & Engineering*, vol. 19, no. 6, pp. 72–78, 2017.
- [13] S. Y. Lee, "UML sequence diagram for reusability of data components," *Journal of System and Management Sciences*, vol. 9, no. 3, pp. 65–77, 2019.
- [14] J. Patchett and J. Ahrens, "Optimizing scientist time through in situ visualization and analysis," *IEEE Computer Graphics and Applications*, vol. 38, no. 1, pp. 119–127, 2018.
- [15] D. S. . Gordeev, "A survey of visualization techniques of algorithms on graphs," *Scientific Visualization*, vol. 10, no. 1, pp. 18–48, 2018.
- [16] M. Pirdehi, "Cloud computing and software services: theory and techniques," *Computing Reviews*, vol. 60, no. 2, pp. 50–51, 2019.
- [17] T. Ogi and T. Yokota, "Effect of visual attention guidance by camera work in visualization using dome display," *International Journal of Modeling Simulation & Scientific Computing*, vol. 9, no. 3, pp. 1840003.1–1840003.13, 2018.
- [18] G. You and S. Leung, "Eulerian based interpolation schemes for flow map construction and line integral computation with applications to Lagrangian coherent structures extraction," *Journal of Scientific Computing*, vol. 74, no. 1, pp. 70–96, 2018.
- [19] M. Xia, W. Liu, B. Shi, L. Weng, and J. Liu, "Cloud/snow recognition for multispectral satellite imagery based on a multidimensional deep residual network," *International Journal of Remote Sensing*, vol. 40, no. 1-2, pp. 156–170, 2019.
- [20] X. Tian, Q. Song, and F. Tian, "Multidimensional data aggregation scheme for smart Grid with differential privacy," *International Journal on Network Security*, vol. 20, no. 6, pp. 1137–1148, 2018.
- [21] P. Liu and Y. Song, "Segmentation of sonar imagery using convolutional neural networks and Markov random field," *Multidimensional Systems and Signal Processing*, vol. 31, no. 1, pp. 1–27, 2020.
- [22] H. Tavaafoghi and D. Teneketzis, "Multidimensional forward contracts under uncertainty for electricity markets," *IEEE Transactions on Control of Network Systems*, vol. 4, no. 3, pp. 511–522, 2017.

## Research Article

# Empirical Study of Large-Scale HLA Simulation of Parallel Region-Matching Knowledge Recognition Algorithm Based on Region Matching

Guohua Zhu  and Haizhou Wang

*School of Artificial Intelligence, Jiangnan University, Wuhan 430056, Hubei, China*

Correspondence should be addressed to Guohua Zhu; zhugh@jhun.edu.cn

Received 28 January 2022; Accepted 24 February 2022; Published 11 April 2022

Academic Editor: Daqing Gong

Copyright © 2022 Guohua Zhu and Haizhou Wang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Most of the existing region-matching algorithms need to match all regions, resulting in a waste of computing resources, increasing the cost of simulation technology and data redundancy, and resulting in the reduction of network data stream transmission efficiency. This paper presents a parallel region-matching knowledge recognition algorithm. Combined with the shortcomings of existing matching algorithms, a simulation technology is constructed to realize the parallel matching of multiple regions in HLA distributed simulation. The algorithm can realize the parallel matching calculation of multiple changed regions in one simulation. At the same time, the basic idea based on mobile intersection is adopted in the matching calculation, and the historical information before and after the region range is moved is used. The matching is limited to the moving interval, and the moving crossover theory is applied to the matching calculation to realize the relevant historical information before and after the region. Simulation results show that the parallel region-matching knowledge recognition algorithm can support HLA distributed simulation evaluation. In the matching calculation, the basic idea based on moving intersection is adopted, and the matching is limited to the moving interval by using the historical information before and after the region is moved, which reduces a large number of irrelevant calculations. Theoretical analysis and experimental results show that the algorithm is particularly suitable for the application needs of building large-scale distributed simulation based on multi-core computing platform.

## 1. Introduction

With the continuous development of social economy, engineering construction projects have started to be implemented one after another [1, 2]. For large-scale construction, relying only on traditional design drawings may not fully meet the needs [3–5]. Therefore, experts within the industry have introduced computer simulation technology to further shorten the corresponding period, improve the corresponding training quality, and save the corresponding expenditure. It has been widely applied in many fields [6–8]. There are also many in-depth explorations in China and other countries, such as proposing corresponding high-level architecture (HLA), creating corresponding distributed simulation based on a flexible and customized architecture, and realizing the interoperability and reusability of various

modules and simulation bodies. For HLA, it mainly includes rules, interface specifications, and object templates. In the specific simulation process, distributed simulation can be regarded as a complete dataset, and each simulation component can be called a member [9, 10]. In the existing HLA simulation process, if the relationship between the simulation models is coupled, it will cause the corresponding system to lack a certain degree of flexibility and cause the overall simulation efficiency to decrease [11, 12]. Therefore, if the dynamic and orderly allocation of resources is to be achieved, the simulated client terminal and the server need to be effectively separated. Some scholars integrate grid technology and simulation technology to improve the above shortcomings and use the separation of the simulated client terminal and the server to achieve dynamic adjustment [12, 13]. Such a fusion method has certain advantages, such

as dynamic adjustment and automatic allocation of data resources, full-process support for the entire life cycle of simulation, strong support for security performance, adaptability to dynamic changes of the grid, automation of resource selection, performability of operating data, automatic collection of simulation results, enhancement of fault tolerance performance, and so on [14, 15].

As the scale of simulation continues to increase, the efficiency of simulation calculation has declined. For simulation entities with a huge amount of data, they all have considerable communication or connection with each other, and these data volumes often show exponential changes or growth. Especially in a specific practical application environment, how to achieve effective data distribution and data simulation in such a situation is a problem worthy of in-depth study [16, 17]. The first is to filter a large amount of data based on related data filtering methods, aiming to reduce the transmission and reception of redundant data during the data simulation operation, thereby reducing the feasibility and data flow of the unit to receive the corresponding redundant data [18–20]. Secondly, as the number of simulation entities increases, HLA distributed simulation experiments will have a large amount of data movement during each simulation experiment [21, 22]. In order to ensure the effectiveness and rationality of data distribution, it is necessary to perform reasonable overlap matching on the overlapping relational areas. Meanwhile, the more the number of entities in the simulation entity, the greater the difficulty of matching. Therefore, the matching algorithm directly determines the efficiency of the simulation and also affects the scalability of corresponding simulation. In view of these needs and deficiencies, based on the parallel region-matching knowledge recognition algorithm, the corresponding simulations are used to achieve parallel matching of multiple regions through combining the business logics of HLA distributed simulation, and the theory of mobile intersection is used in matching calculations to achieve the historical information related to forward or backtracking region, which is limited to a fixed threshold, reducing invalid operations and aiming to improve the effectiveness of HLA distributed simulation.

### 1.1. HLA Distributed Evaluation Method

*1.1.1. Concepts Related to Regional Matching.* In view of area matching, in the specific simulation specification, the management of data distribution is based on the corresponding area. The data collector or producer uses the effective designation of the fixed area to achieve data production, and the data user uses search for a specific area to achieve reception of the data. Take a specific area overlap as an example, as shown in Figure 1.

*Definition 1.* Dimension: the so-called dimension is based on a specific named interval, which is essentially a non-negative interval. The downward limit is 0, and the upward limit varies according to the specific dimensions. Its value needs to be defined in the corresponding dimension table.

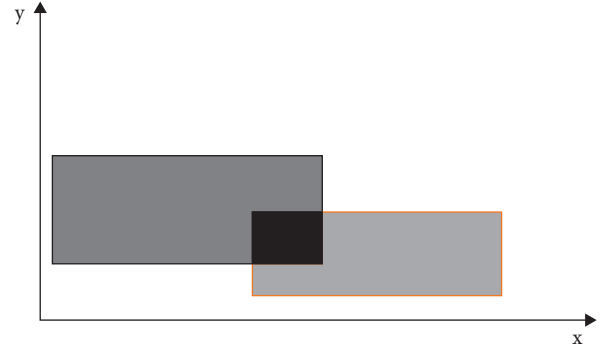


FIGURE 1: Description of overlapping regions in a two-dimensional interest space.

*Definition 2.* Range: it is different from dimension. It is a continuous integer half-open interval and is a sub-dataset of related dimensions. It is specifically defined by an ordered integer. The specific first number can be considered as the downward limit of the range, and the second integer is the upward limit.

*Definition 3.* Region description: it is a specific range dataset, that is, the corresponding dimension described by the range contained in the area is used. For each dimension described in a specific area, it can only have one range.

*Definition 4.* Region realization: the so-called region realization is to realize attribute update, interactivity, data query, etc. through specific associations.

*Definition 5.* Region: it is the general term for the description and realization of the region.

*Definition 6.* Update region (publish region): The parallel area mainly realizes the real-time update of data to meet the conditions of data query.

*1.1.2. Parallel Region-Matching Knowledge Recognition Algorithm.* The essence of the so-called region-matching algorithm is to determine whether there is a corresponding overlap in the individual interest units of the region, and the specific matching efficiency and accuracy determine the relevant efficiency of specific interest matching [23, 24]. The specific typical matching algorithm can be divided into direct matching, grid matching, mixed matching, classification matching, mobile intersection matching, etc. according to the matching principle or method [25, 26].

For direct matching, its essence is to use a separate data structure to update the range of data and the specific range of query subscriptions. When users query specific data, the algorithm in this paper will identify the corresponding data information according to specific conditions and then calculate the overlapping area with the identification. This method is relatively simple, not requiring additional information, and can ensure accurate matching. However, this algorithm has a relatively large computational complexity. For a large number of areas of large-scale simulation, this

algorithm is difficult to adapt to this simulation calculation [27, 28].

For grid matching, the essence is to use the simulation system to divide the full-dimensional regular grid in advance and set the corresponding data-related channel for each network logic, especially for the specific calculation of the regional boundary to determine the correlated network covered by the region, so as to clarify the relevant channels of data transmission. If the data update area and the query data area have the same unit, then grid matching can be performed. The complexity of the network matching algorithm is relatively simple compared with direct matching, but network matching often generates redundant links and requires manual data filtering. From the overall effect, the specific network matching algorithm is difficult to reconcile in the specific matching speed and accuracy to a certain extent, and the size of the grid affects the efficiency of the overall algorithm.

In regard to the limitations of the direct matching and network matching algorithms, some scholars have proposed a related algorithm of hybrid matching, which uses specific network matching to achieve one-to-one mapping of all regions to a specific network and realizes the exact matching simple compared with specific network. On the one hand, this hybrid algorithm can ensure the accuracy of matching but also reduce the complexity of the entire direct matching algorithm. However, consistent with the network algorithm, the integrity of the hybrid matching algorithm still needs to be restricted by the size of the network.

For the classification matching algorithm, its essence is to use all areas of the system to perform multi-dimensional projection sorting, realize the overlap analysis of two or more dimensions on the projection, and indicate that these ranges have a certain intersection and overlap. The complexity of the classification matching algorithm is relatively moderate, that is, the classification matching algorithm can be used to complete the overlap judgment of the corresponding area at one time. This judgment can realize the misalignment of the operation and the improvement of the matching efficiency. However, because parallel matching recognition is to effectively detect the corresponding region, in the actual simulation process, if some recognition regions change, the system will be recalculated, resulting in the decline of recognition rate.

The specific mobile matching is used to achieve dynamic matching, the essence of which is that only part of the region of the large-scale simulation system changes, so it only needs to change according to the information before and after the area change to realize the information movement of the matching calculation and the effective change of matching range, while effectively reducing the amount of specific matching calculations. During the process of mobile matching, the specific algorithm matching degree only needs to be related to the number of change regions and has nothing to do with the total number of regions of the entire system. Therefore, it can be seen that mobile matching is not only dynamic but also accurate and efficient.

Therefore, it is necessary to use multi-core CPUs for actual processing. When the computing platform becomes multi-core and uses distributed interactive simulation on a larger scale, specific high-performance parallel computing can be used to improve specific simulation performance and achieve support for complex systems. Although the above matching algorithm causes certain restrictions and limitations to large-scale simulation calculations, there are also certain limitations, which are restricted by the bottleneck of computational complexity. Therefore, it is worthy of in-depth exploration and research to realize efficient parallel computing on a multi-core, distributed computing platform.

First, a specific domain scanning method is used to collect data, and the specific sequence is shown in Figure 2. When there is overlap between  $p$  and the surrounding area pixels, it is necessary to fully consider the scanning speed to avoid repetitive scanning. Figure 2(b) shows a specific scanning scheme, in which the shaded parts of the diagonal lines are overlapping areas. During scanning, these parts do not need to be scanned repeatedly, and only important grid parts are scanned for specific pixel point judgment. [1, 8] is used to determine the scan template.

A specific two-dimensional array is used as a scanning template, each specific row is used as a scanning template, each column of data is used as the number of scanning points included, and the rest are filled with 0. When a specific pixel point is scanned, the adjacent points need to be assigned to the specific periphery, and the circular labeling is realized according to the specific point.

The breadth-first method is used to search the parallel region-matching knowledge recognition algorithm. Assuming that the image is a two-dimensional image with  $M$  rows and  $N$  columns, the valueless represents black pixels, and value 1 represents white pixels. The specific steps are shown below.

- (1) First, create a specific tag array to indicate whether the pixels of the image have been processed. When the image is scanned, the array with tag attributes gives different tag values for each specific area.
- (2) The principle of left-to-right and top-to-bottom is used to effectively scan the image. When realizing a specific marked pixel scan, if the pixel is marked, it needs to be recorded for subsequent mark analysis.
- (3) When searching for the area at the starting point of the area, a fixed algorithm needs to be specifically called, and the label value of the area is set to a fixed value; when the function completes the specific call, the coordinates of all pixels in the area can be obtained.
- (4) Scan the image through  $p$ , repeat steps 2 and 3 iteratively, until the specific image is scanned, and the algorithm process ends.

On the basis of the above steps, the specific number of regions can be obtained, and the corresponding parameters can be used to mark each part of the region to achieve the specific mark value to obtain the corresponding region containing points. In the specific steps, it is necessary to



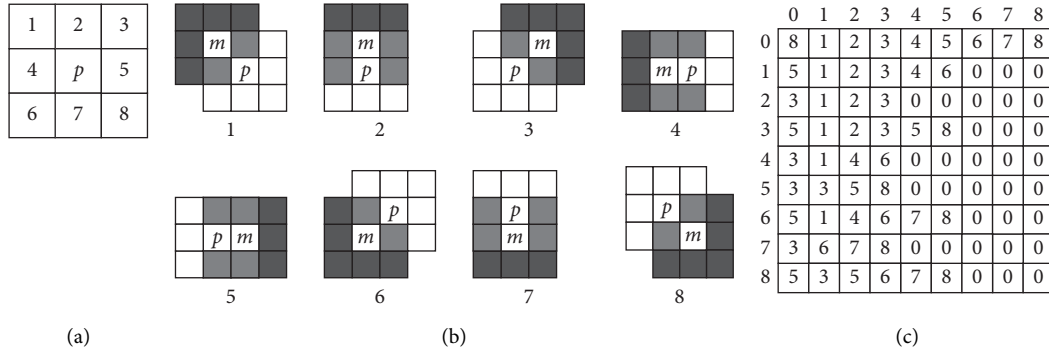


FIGURE 2: Pixel 8 neighborhood priority order, scan template, and two-dimensional array. (a) Scan-line order. (b) Scanning templates for *p*'s different neighborhoods. (c) Array of the scanning templates.

implement the algorithm call of the connected region through the algorithm, and the single-region binary graph is shown in Figure 3.

In order to realize the flow of the algorithm, the serial number is marked according to the specific search order. As shown in Figure 4, when the image is scanned to a specific value (1), the value of the corresponding array is recorded as 0, and it is taken as the starting point and ending point of the new region, the total number of corresponding regions is increased accordingly, and in the meantime, the value of the array is recorded as the corresponding array. On the one hand, it can indicate that this point has been scanned accordingly; on the other hand, it also belongs to certain region.

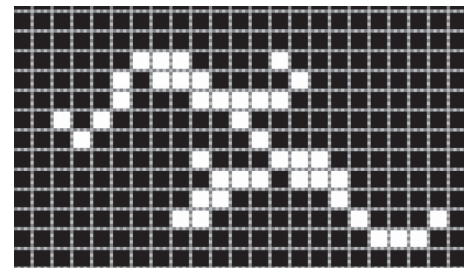


FIGURE 3: Binary graph of a single region.

- (1) Use the corresponding queue to record the coordinates of the corresponding recording point and the scan mark.
- (2) The circular identification is performed for the corresponding queue; it is necessary to ensure that the queue is not null. When reading the point value at the head of the queue, it needs to be identified according to the specific scan template, and the inspection and marking are given in the corresponding order. Meanwhile, the marked points that are not scanned are appended to the queue through the corresponding array value, and then the scanned point is removed.

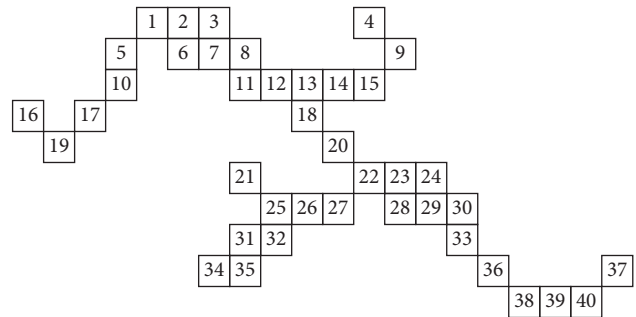


FIGURE 4: Pixel numbers given from top left to bottom right.

The specific calling process is shown in Figure 5, the number in the circle represents the specific label of the scan template at a certain point, and the specific underlined point can be identified as the critical point of the same point. It can be seen from the running results of Figure 5 that if a different scanning sequence is used to scan the region of pixels, it is often unnecessary to traverse the surrounding points, and only 5 points need to be traversed at most. This scanning method effectively improves the operating efficiency. There is no specific mark conflict in the whole scanning process, and the final region will be scanned in the corresponding order from top left to bottom right.

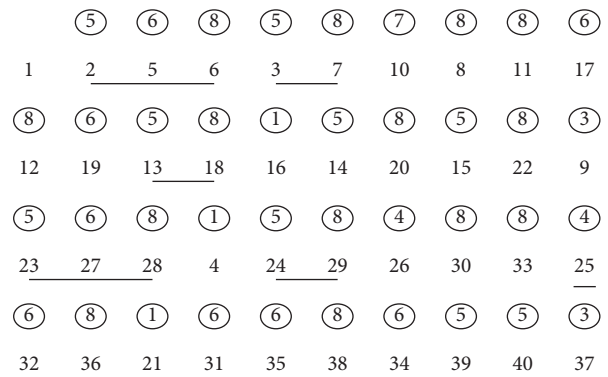


FIGURE 5: Points of the regional operation process.

As the computing performance of the computer has been significantly improved, in order to facilitate the use of the computer, make maximum use of computing power, use

connectivity detection to analyze the effectiveness of each pixel, and analyze the specific connection relation between the pixels before and after. The specific region merging is

realized through related parallel algorithms, and different algorithms are used to realize the different complexity of parallel algorithms in dealing with specific region overlaps. If a specific function is called directly, you can directly obtain the relevant computing performance of the computer and divide the image according to the computing power. If it is a two-core computing power, it is divided into two parts, and the region detection of the two parts is performed according to the above method. When the calculation is completed, the overlapping region is effectively scanned, and then the relevant regions are merged.

By setting the corresponding image size, using multiple computers for calculation, and using specific parameters to

represent specific threads, formulas (1) and (2) are used to calculate the start lines and end lines of the thread.

$$starline = (threadNum - 1) \times H/N, \quad (1)$$

$$endline = threadNum \times H/N - 1. \quad (2)$$

HLA-distributed simulation data are collected by sensor equipment and stored in large capacity data storage equipment. After data processing, it is sent to the corresponding program for operation.

$$T(a, a_1) = \frac{\sum_{i=1}^n ((q_i - s)^2 (q_i - s - l))(q_i - s^2 (q_i - s - l))}{\sqrt{\sum_{i=1}^n (q_i - s^2 (q_i - s - l))^2 + (q_i - s^2 (q_i - s - l))^2}}, \quad (3)$$

where  $T(a, a_1)$  is the set of characteristic attributes of HLA distributed data and HLA distributed characteristic expression;  $q_i$  is the number of data features after classification for HLA distributed data; and  $s$  is the feature content of the HLA distribution. It is a numeric argument, a property specific to HLA distributed data. After identifying the characteristics of HLA distributed data, non-characteristic attributes must be removed. It can reduce the error and improve the speed when collecting. The redundant data removal formula is as follows:

$$L = \vec{q} + \frac{\sum_{i=1}^n (T(a, a_1))_i (q - e)}{\sum_{i=1}^n [T(a, a_1)]_i^2}, \quad (4)$$

where  $L$  defines the removal benchmark and removes those that meet the benchmark;  $\vec{q}$  indicates the filter request to be used when removing;  $e$  represents an existing redundant data removal request. The features of the data can be obtained by filtering.

$$C = \beta \left( \prod_{i=1}^n e_i + \frac{\sum_{i=1}^n T(a, a_1) \cdot (a, a_1)_i d^{i\theta}}{\sum_{i=1}^n T(a, a_1)_i^2} \right) + (1 - \beta) \left( q_i + \frac{\sum_{i=1}^n (T(a, a_1)_i - \bar{X}) \cdot T(a, a_1)_i^2}{\sum_{i=1}^n (T(a, a_1) - \bar{X})^2} \right), \quad (5)$$

where  $\alpha$  represents the ownership value of the data feature and  $\beta$  represents the correlation coefficient of the balance factor. Assume input vectors  $x_i$  and label values  $y_i$ . The softmax loss function expression used in this paper is

$$L = -\frac{1}{m} \sum_{i=1}^m \log \frac{e^{f_{y_i}}}{\sum_{j=1}^n e^{f_j}} = -\frac{1}{m} \sum_{i=1}^m \log \frac{e^{W_{y_i}^T x_i + b_{y_i}}}{\sum_{j=1}^n e^{W_j^T x_i + b_j}}, \quad (6)$$

where  $x_i$  represents the  $i^{th}$  HLA distributed simulation evaluation efficacy image feature;  $y_i$  represents the real category label corresponding to  $i^{th}$  HLA distributed simulation evaluation efficacy image;  $W_j$  represents the category

weight;  $b_j$  represents the error value of the category;  $m$  and  $n$  represent the number of training samples and the number of categories in turn; and  $f_j$  represents the inner product relationship  $f_j = W_j^T x_i + b_j$  between the category weight  $W_j$  and the bias value  $b_j$  when the fully connected layer is activated. The HLA distributed simulation evaluation efficacy identification technology should satisfy the following conditions in the HLA distributed simulation evaluation efficacy feature value: the distance between the same HLA distributed simulation evaluation efficacy features needs to be minimized, and the distance between different HLA distributed simulation evaluation efficacy features needs to be maximized.

The corresponding sequential marking algorithm is used to mark the overlapping regions, which cannot ensure the sequential arrangement, so the process of merging the regions will be more complicated. Therefore, this paper proposes a reverse merging algorithm for this limitation. That is, in the process of reverse merging, the processing is performed from the second-to-last overlapping row, and each time a row is processed; the corresponding thread will analyze and correspond to determine whether the merging process is required.

In view of the limitations of existing region-matching algorithms, this paper constructs a parallel region-matching algorithm for HLA distributed simulation, which integrates mobile matching and parallel computing methods and divides the tasks of multiple mobile matching in the simulation into different cores, and the multi-threading methods are used to achieve specific parallelized calculation of region matching, so as to improve the calculation of region-matching ability. Similarly, in each specific thread calculation, the method of using a mobile region to match the region is realized, to reduce the number of redundant calculations and improve the actual efficiency of matching.

For the parallel region-matching knowledge recognition algorithm, its specific principles are as follows. First, when a specific single region is effectively moved, the overlap change between other regions and the part of the region is actually

related to the movement of the single region. Therefore, when updating a specific region, there is no need to consider the range outside the actual region movement region. The overlap of these regions cannot be directly changed; in the specific simulation process, there is certain irrelevance in regions of the same type. The changes in multiple regions of the same type are overlapped by multiple calculations, as shown in Figure 6.

Figure 6 shows the changes in the overlap of the two-dimensional region during the specific simulation advancement process. When a specific change occurs in the corresponding region, a specific update region can be realized through the boundary of the region, and specific overlapping parallel calculations can be realized from the specific region, and the parallel computing performance can be effectively realized.

When a specific region changes to a certain extent, the relevant core of the matching calculation is caused according to the region, and the position query before and after the movement is realized to ensure the region where the region moves. Here, the dimensional data stored in the region are realized in an orderly manner through a specific index, that is, the data of each dimension in a specific multi-dimensional space are used along with the index for storage. On the one hand, it uses the indexed ordered table to store the specific update region data. On the other hand, it is used to store the range boundary value of various region projections in specific dimensions. Each group is combined through two specific ordered tables, one index table stores the closed point of lower boundary of the range, and the other index table is used to store the opening point of the upper boundary. The definition of the specific data structure of the node is as follows:

```

struct Node {
    int id; //range id
    Struct Node * Next; //Node pointer
}

```

The ordered list of each index implements a specific pointer array index, uses the size of the array to set the dimension, and compares the boundary value with the corresponding index value of the array element, as shown in Figure 7.

It can be seen from Figure 7 that for fixed nodes with no moving range, redundant operation analysis that is not within the operating range can be achieved without effective matching, which effectively saves related computing resources and can achieve accurate analysis and matching in specific ranges, and there is no specific false connection problem.

From the framework of Figure 8, in view of multi-threading, the calculation of overlapping regions is realized, including control threads and multiple computing threads, to complete the calculation task through the task queue. After a specific movement range occurs in the control thread region, it is responsible for the specific task generation node, and it is listed in the specific task queue, and the calculation of the result is completed by the calculation

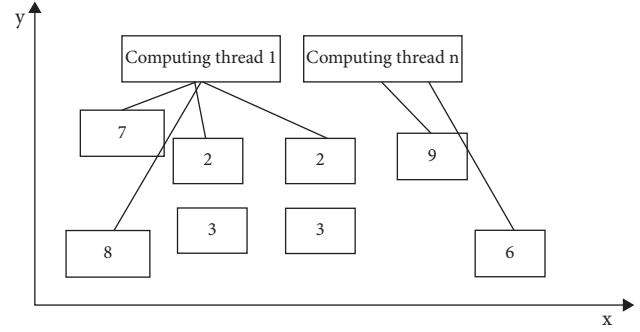


FIGURE 6: Schematic diagram of the intersecting changes of regional movement.

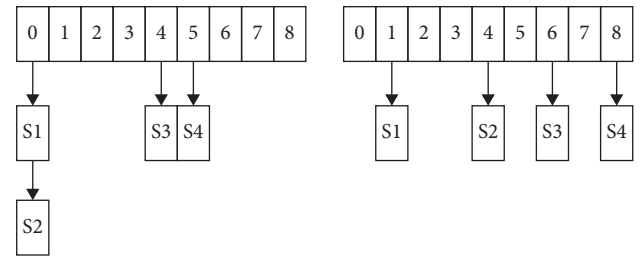


FIGURE 7: Range representation based on indexed ordered list.

of the control thread. In each simulation advancement process, the first thing that each computing thread needs to process is the update region, until all tasks are processed. The semaphore mechanism is used to synchronize between the threads.

The specific steps of the so-called control thread algorithm are as follows.

For (every simulation advancement process)

- (1) For each announcement range  $a_m$ , if  $a_m$  moves, insert  $a_m$  into the announcement task queue pq.
  - (2) For each order range  $b_n$ , if  $b_n$  moves, insert  $b_n$  into the order task queue sq.
  - (3) If pq is not null, for each computing thread 1, release the semaphore  $m\_publish\_i$ , which is used to notify the computing thread to start processing and update the task queue.
  - (4) Wait for the completion semaphore  $m\_publish\_finish\_i$  of each computing thread  $i$  until all update tasks are processed.
  - (5) If sq is not null, for each calculation line  $i$ , release the semaphore  $m\_subscribe\_i$ , which is used to notify the calculation thread to start processing the order task queue.
  - (6) Wait for the completion semaphore  $m\_subscribe\_finish\_i$  of each calculation thread  $i$  until all the subscription overlap calculations are completed.
- end for

**1.1.3. Simulation Experiment.** In order to verify the effectiveness of the parallel region-matching knowledge

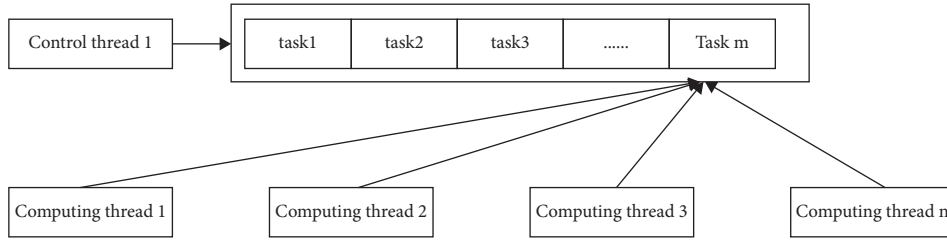


FIGURE 8: Parallel region-matching method framework.

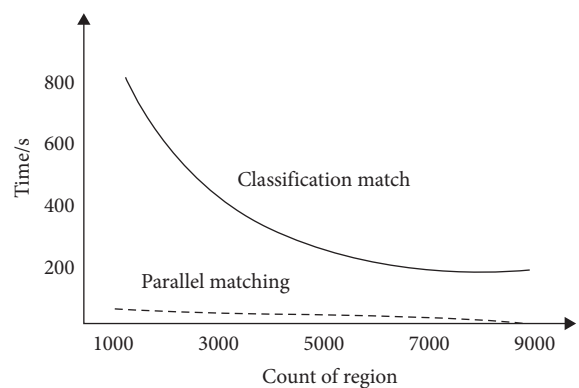
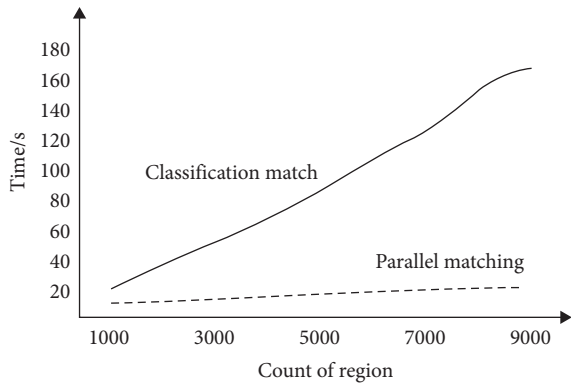
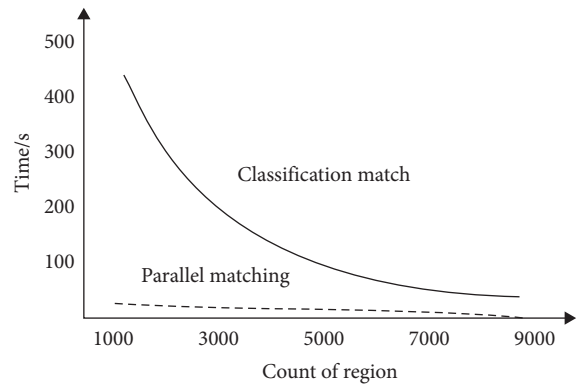
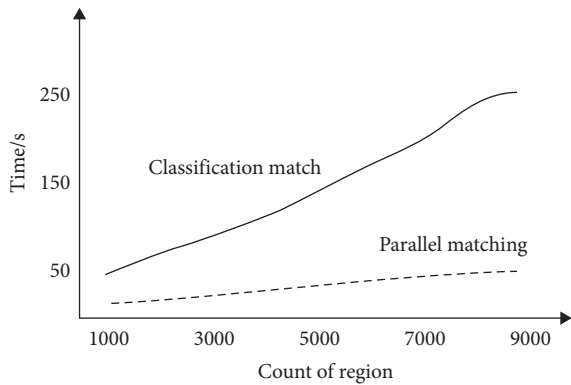
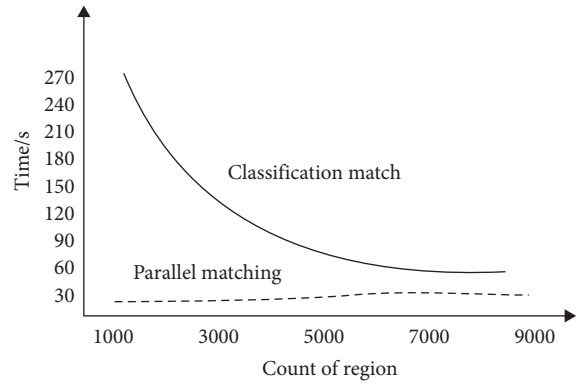
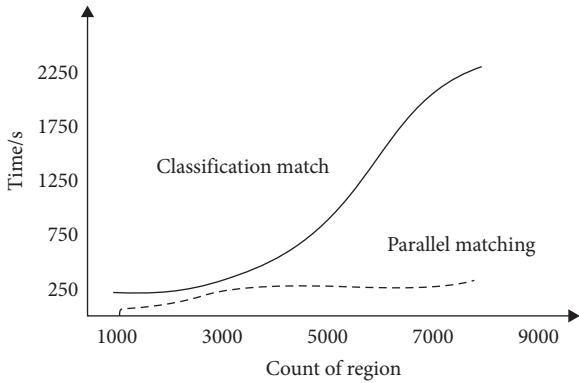


FIGURE 9: Algorithm performance test results for different numbers of regions (2000, 4000, 8000).

FIGURE 10: Algorithm performance test results at different dimensional upper limits.

recognition algorithm, the following simulation experiments are set up in this paper. First, a specific two-dimensional space is generated according to specific user input, and a specific publishing region and ordering region are generated

in the two-dimensional space, and corresponding algorithm comparison experiments are designed.

Corresponding tests are carried out for the algorithm performance in different numbers of regions, and the HLA

distributed simulation evaluation changes of the matching algorithm in the corresponding cases of 2000, 4000, and 8000 are calculated, respectively, as shown in Figure 9. It can be seen from the results that the parallel region-matching knowledge recognition algorithm is relatively appropriate in terms of time consumption.

For different upper dimensional limits, as shown in Figure 10, the performance of parallel region-matching knowledge recognition algorithm is obviously better than other algorithms. When the upper limit of dimensionality is greater than 5000, the algorithm performance is the same as the grid matching performance with a grid number of 100, and there is no false connection phenomenon in grid matching.

The parallel region-matching knowledge recognition algorithm is closely related to the number of threads of calculation. From the corresponding simulation experiment results, compared with the traditional method, the parallel region-matching knowledge recognition algorithm does not change the performance with the increase of the specific number of regions, but in terms of specific dimensions, it has a change relationship with the specific upper limit of the dimension. Through multi-core computing power, better acceleration ratio is realized, especially for large-scale regional changes. Therefore, the parallel region-matching knowledge recognition algorithm is more suitable for HLA distributed simulation evaluation.

## 2. Conclusion

The development of industry requires HLA distributed simulation to provide specific support for the long-term, stable, and healthy development of industrial economy. Most of the existing region-matching algorithms need to match all regions, resulting in a waste of computing resources. At the same time, it is difficult to give full play to the parallel computing advantages of multi-core platform mainly based on the idea of serial matching. In view of these needs and shortcomings, based on the parallel region-matching knowledge recognition algorithm, this paper realizes the parallelism of region-matching calculation through multi-core platform. At the same time, for HLA distributed simulation, this paper constructs a simulation system to realize multi-region parallel matching. In the matching calculation, the mobile crossover theory is used to realize the relevant historical information of regional prepositions or backtracking, and it is limited to a fixed threshold range to reduce invalid operations. It has efficient matching and good acceleration performance and can support HLA distributed simulation evaluation. The experimental results show that the algorithm has high matching efficiency, does not cause false connections, and has good acceleration performance. It can give full play to the computing performance of multi-core computing platform and meet the needs of large-scale distributed simulation data distribution and management.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References

- [1] J. Sauter, U. V. Solloch, A. S. Giani, A. J. Hofmann, and A. H. Schmidt, "Simulation shows that HLA-matched stem cell donors can remain unidentified in donor searches[J]," *Scientific Reports*, vol. 6, no. 1, pp. 119–129, 2016.
- [2] Y. Park and D. Min, "HLA-DDS API transformation for DDS communication based HLA simulation," *Advanced Science Letters*, vol. 21, no. 3, pp. 244–252, 2015.
- [3] Y. Bouteraa, I. B. Abdallah, A. ElMogy, A. Ibrahim, U. Tariq, and T. Ahmad, "A fuzzy logic architecture for rehabilitation robotic systems," *International Journal of Computers, Communications & Control*, vol. 15, no. 4, pp. 1–17, 2020.
- [4] A. Falcone, A. Garro, S. . Taylor, and A. . . Anagnostou, "Experiences in simplifying distributed simulation: the HLA development kit framework," *Journal of Simulation*, vol. 11, no. 3, pp. 208–227, 2017.
- [5] S. Jung-Yong and H. Soumg, "An implementation of HLA standard for weapon system simulation[J]," *Communications of the Korean Institute of Information Scientists & Engineers*, vol. 86, no. 3, pp. 309–321, 2016.
- [6] X. Feng and J. Gao, "Gene sequences parallel alignment model based on multiple inputs and outputs," *International Journal of Computers, Communications & Control*, vol. 14, no. 2, pp. 141–153, 2019.
- [7] Y. Wu and G. Gong, "A real-time scheduling algorithm for HLA-based simulation models[J]," *International Journal of Modeling Simulation and Scientific Computing*, vol. 22, no. 1, pp. 95–96, 2015.
- [8] Gorecki, Bouanan, and Zacharewicz, "Integrating HLA-based distributed simulation for management science and BPMN," *IFAC-PapersOnLine*, vol. 51, no. 11, pp. 655–660, 2018.
- [9] Q. Wang, A. I Li-Rong, and A. Z. Gong, "Study on load balancing in hla-based distributed simulation system," *Computer Technology and Development*, vol. 3, no. 2, pp. 16–19, 2019.
- [10] C. Roth, O. Sander, M. Kühnle, and J. Becker, "Hla-based simulation environment for distributed systemc simulation," *Ai Trade News*, vol. 5, no. 11, pp. 89–94, 2013.
- [11] D. Guo, J. Zhang, and Q. Liang, "Design of distributed training simulation system based on hla," *Radio Engineering*, vol. 70, no. 1, pp. 1–9, 2019.
- [12] C. Sato, K. Mio, and T. Ogura, "3P-031 molecular dynamics simulation studies on the differences in the binding mechanism of LILRB1/HLA-G and LILRB2/HLA-G(Protein: Structure & Function,The 47th annual meeting of the biophysical society of Japan)[J]," *Seibutsu Butsuri*, vol. 49, no. 5, pp. 10–20, 2017.
- [13] A. Nowroozi and M. Shahlaei, "A coupling of homology modeling with multiple molecular dynamics simulation for identifying representative conformation of GPCR structures: a case study on human bombesin receptor subtype-3," *Journal of Biomolecular Structure and Dynamics*, vol. 35, no. 2, pp. 250–272, 2016.
- [14] J. Kwok, M. Guo, W. Yang et al., "Simulation of non-inherited maternal antigens acceptable HLA mismatches to increase the chance of matched cord blood units: Hong Kong's experience," *Human Immunology*, vol. 3, no. 4, pp. 113–120, 2018.
- [15] Y. You, L. Tan, T. Lee, W. Kim, and S. Yoon, "Development of an OMT table viewer/editor using the ms-based distributed

- simulation,” *International Journal of Information and Electronics Engineering*, vol. 6, no. 2, pp. 89–92, 2016.
- [16] A. . Albagli, D. . Falcão, and J. . de Rezende, “Smart grid framework co-simulation using HLA architecture,” *Electric Power Systems Research*, vol. 130, no. 4, pp. 22–33, 2016.
- [17] S Tafulo, J Malheiro, L Dias et al., “Improving HLA matching in living donor kidney transplantation using kidney paired exchange program,” *Transplant Immunology*, vol. 62, no. 2, pp. 101–117, 2020.
- [18] S. Lee, S. Lee, K. Hwang, and S. Kim, “Design and implementation of the multi-resolution interoperation simulation using HLA/RTI,” *Journal of the Korea Society for Simulation*, vol. 24, no. 1, pp. 9–16, 2015.
- [19] P. Ozbek, “Dynamic characterization of HLA-B\*44 Alleles: a comparative molecular dynamics simulation study,” *Computational Biology and Chemistry*, vol. 62, no. 2, pp. 12–16, 2016.
- [20] J. Kwok, W. H. Tang, W. K. Chu et al., “High resolution allele genotyping and haplotype frequencies for NGS based HLA 11 loci of 5266 Hong Kong Chinese bone marrow donors[[]],” *Human Immunology*, vol. 81, no. 10, pp. 1–9, 2020.
- [21] F. R. Rad, M. G. Akbari, M. Zamani, S. Bayat, and M. Zamani, “Pharmacogenetic and association studies on the influence of HLA alleles and rivastigmine on the Iranian patients with late-onset alzheimer’s disease,” *Molecular Neurobiology*, vol. 2, no. 1, pp. 10–20, 2021.
- [22] B. H. Shin, M. J. Everly, and H. Zhang, “Impact of tocilizumab (Anti-IL-6R) treatment on immunoglobulins and anti-HLA antibodies in kidney transplant patients with chronic antibody-mediated rejection[[]],” *Transplantation*, vol. 104, no. 2, pp. 109–118, 2020.
- [23] D. Chatterjee, P. Priyadarshini, D. . Das, K.. Mushtaq, and J. . Agrewala, “Deciphering the structural enigma of HLA class-II binding peptides for enhanced ipve,” *Journal of Proteome Research*, vol. 19, no. 11, pp. 4655–4669, 2020.
- [24] M. Arévalo, C. López-Medina, M. M. Martínez-Losa et al., “Role of HLA-B27 in the comorbidities observed in Axial Spondyloarthritis: data from COMOSPA,” *Joint Bone Spine*, vol. 87, no. 5, pp. 1–9, 2020.
- [25] A. Hernández-Hernández, B. Hernández-Zaragoza, R. Barquera et al., “Genetic diversity of HLA system in two populations from Oaxaca, Mexico: oaxaca city and rural Oaxaca - ScienceDirect,” *Human Immunology*, vol. 81, no. 9, pp. 553–556, 2020.
- [26] Y. Liu, H. Sun, W. Fan, and T. Xiao, “A parallel matching algorithm based on order relation for HLA data distribution management,” *International Journal of Modeling Simulation and Scientific Computing*, vol. 06, no. 2, pp. 73–78, 2015.
- [27] A. Abed, L. Calapre, J. Lo et al., “301MO Genomic HLA as a predictive biomarker for survival among non-small cell lung cancer patient treated with single agent immunotherapy,” *Annals of Oncology*, vol. 31, no. 5, pp. 1359–1366, 2020.
- [28] E. Fainardi, D. Bortolotti, M. Castellazzi, I. Casetta, T. Bellini, and R. Rizzo, “Detection of serum soluble HLA-G levels in patients with acute ischemic stroke: a pilot study - ScienceDirect,” *Human Immunology*, vol. 81, no. 4, pp. 156–161, 2020.



## Retraction

# Retracted: Dynamic Performance Analysis of STEP System in Internet of Vehicles Based on Queuing Theory

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] H. Li, X. Wu, and Y. Wang, "Dynamic Performance Analysis of STEP System in Internet of Vehicles Based on Queuing Theory," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 8322029, 13 pages, 2022.

## Research Article

# Dynamic Performance Analysis of STEP System in Internet of Vehicles Based on Queuing Theory

Huiyong Li , Xiaofeng Wu , and Yanhong Wang 

*School of Intelligent Manufacturing and Information, Jiangsu Shipping College, Nantong 226010, China*

Correspondence should be addressed to Huiyong Li; [lihuiyong@jssc.edu.cn](mailto:lihuiyong@jssc.edu.cn)

Received 29 January 2022; Revised 4 March 2022; Accepted 9 March 2022; Published 10 April 2022

Academic Editor: Daqing Gong

Copyright © 2022 Huiyong Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The Internet of vehicles (IoV) is an important research area of the intelligent transportation systems using Internet of things theory. The complex event processing technology is a basic issue for processing the data stream in IoV. In recent years, many researchers process the temporal and spatial data flow by complex event processing technology. Spatial Temporal Event Processing (STEP) is a complex event query language focusing on the temporal and spatial data flow in Internet of vehicles. There are four processing models of the event stream processing system based on the complex event query language: finite automata model, matching tree model, directed acyclic graph model, and Petri net model. In addition, the worst-case response time of the event stream processing system is an important indicator of evaluating the performance of the system. Firstly, this paper proposed a core algorithm of the temporal and spatial event stream processing program based on STEP by Petri net model. Secondly, we proposed a novel method to estimate the worst-case response time of the event stream processing system, which is based on stochastic Petri net and queuing theory. Finally, through the simulation experiment based on queuing theory, this paper proves that the data stream processing system based on STEP has good dynamic performance in processing the spatiotemporal data stream in Internet of vehicles.

## 1. Introduction

Internet of vehicles is an important application of intelligent transportation with Internet of things, and it is also an important part of smart city [1–6]. In Internet of vehicles system, there are various types of data sources that produce large amounts of temporal and spatial uninterrupted data flow. Processing these data stream is an important research direction of big data [7]. Now, there are some products of processing the data flow of Internet: Twitter's Storm system [8], Yahoo's simple scalable streaming system (S4) [9], Facebook's Data Freeway and Puma system [10], LinkedIn's Kafka system [11], and an interactive real-time computing systems Berkeley's Spark [12].

Because the data stream of Internet of vehicles is different from the Internet, therefore, we cannot directly use the Internet data stream processing method to process the temporal and spatial data stream of Internet of vehicles. In recent years, there is some progress in processing the

temporal and spatial data streams of Internet of vehicles using the complex event processing technology. In the event driven architecture Internet of vehicles, all kinds of temporal and spatial data stream generated by the sensing devices were abstracted into basic temporal and spatial events. Through a complex event query language, the system will get the meaningful complex event from the basic temporal and spatial event by certain pattern matching filter. Moody proposed a complex event query language SpaTec, and it will be applied to London's bus monitoring system [13, 14]. Jin et al. have proposed a complex event query language CPSL, and it can describe a variety of spatial and temporal properties of the event [15]. Li and Chen have proposed a temporal and spatial constraints event query language STEP, which can effectively describe the Temporal and Spatial constraint information of Internet of vehicles [16]. Xu et al. have proposed a spatiotemporal event interaction model STEIM to express the spatiotemporal information of the spatiotemporal event flow in the V2X [17].

There are four types processing model of the event stream processing system based on complex event query language: finite automata model, matching tree model, a directed acyclic graph model, and Petri net model. The worst response time of the event stream processing system is an important indicator of the performance evaluation of the system. Akili and Weidlich began to focus on the worst response time of the event stream processing system [18]. Um et al. have proposed an associated constraint labelling method to estimate the worst response time of the event stream processing system [19]. Jiang et al. present a model of the outward transmission of vehicle blockchain data and then give detailed theoretical analysis and numerical results [20]. Hu et al. present a novel double-layered federation model to analyze the performance of the data processing system of Internet of vehicles [21].

The current work has not analyzed the dynamic performance of the stream processing algorithm. In this paper, the dynamic performance of the algorithm is analyzed through the queuing theory model, thereby verifying the feasibility of the algorithm theoretically. In this paper, we proposed a method to estimate the worst case response time

of the STEP event stream processing system, which is based on stochastic Petri net and queuing theory. The simulation results demonstrate the effectiveness of the method. The remainder of this paper is structured as follows: Section 2 introduces the complex event query language STEP. Section 3 introduces the stochastic Petri net theory and queuing theory. Section 4 gives the architecture of the STEP event stream processing system and the core algorithms. Section 5 analyzes four types structured event stream processing program instances by queuing theory. Section 6 verifies the relevant conclusion is correct through the simulation. Finally, the last section concludes this paper.

## 2. Event Query Language: STEP

STEP is a Spatial-Temporal Constraint Event Query Language, which focuses on processing the spatial and temporal constraint data of the Internet of vehicles. The syntax of STEP contains five types of Boolean expressions, and the event instances of Boolean expression EBEXP were composed by the others. Syntax rules of STEP are as follows:

The object identifier Boolean expression OBEXP:

$$obexp ::= \text{TRUE} | \text{FALSE} | x_{id} = ID | x_{id} \neq ID | obexp_1 \wedge obexp_2 | obexp_1 \vee obexp_2. \quad (1)$$

The general numerical attributes Boolean expression ABEXP:

$$abexp ::= \text{TRUE} | \text{FALSE} | x_a = A | x_a \neq A | |x_a > A | x_a < A, \\ | abexp_1 \wedge abexp_2 | abexp_1 \vee abexp_2. \quad (2)$$

The temporal properties Boolean expression TBEXP:

$$tbexp ::= \text{TRUE} | \text{FALSE} | x_t \text{ BEFORE } T | x_t \text{ EQUAL } T | |x_t \\ \text{ OVERLAP } T | x_t \text{ DURING } T, \\ | tbexp_1 \wedge tbexp_2 | tbexp_1 \vee tbexp_2. \quad (3)$$

The spatial properties Boolean expression LBEXP:

$$lbexp ::= \text{TRUE} | \text{FALSE} | x_t \text{ EQ } LOC | x_t \text{ OPLOC} | x_t \text{ IN } LOC | x_t \text{ NORTH } LOC, \\ | lbexp_1 \wedge lbexp_2 | lbexp_1 \vee lbexp_2. \quad (4)$$

The event instances Boolean expression EBEXP:

$$ebexp ::= \text{TRUE} | \text{FALSE} | obexp | tbexp | |lbexp | abexp | ebexp_1 \wedge ebexp_2 | ebexp_1 \vee ebexp_2. \quad (5)$$

## 3. Stochastic Petri Net and Queuing Theory

**3.1. Stochastic Petri Net.** Stochastic Petri net model linked a random delay time with each transition between the case where it can be implemented and its implementation. The performance analysis of stochastic Petri net is built on the foundation that its state space is isomorphism as the Markov chain. Stochastic Petri net provides a good description means of the

performance model of the system, and Markov random process provides a solid mathematical foundation of its evaluation.

*Definition 1.* In continuous time stochastic Petri net, a transition requires a delay time from the case where it can be implemented to the case where it has implemented. It is seen as a continuous random variable  $x_i$  (take a positive real number). Because of the speed of the data flow in the

queuing system subjecting to the Poisson distribution, therefore, the delay time subject to an exponential distribution function is as follows:

$$\forall t \in T: F_t = 1 - e^{-\lambda_t x}. \quad (6)$$

Stochastic Petri net is denoted as  $SPN = (S, T; F, M_0, \lambda)$ , where  $((S, T; F, M_0))$  is a Petri net system, and  $\lambda = \{\lambda_1, \lambda_2, \dots, \lambda_m\}$  is the average implementation rate collection of the transition. Among them, the real parameter  $\lambda t > 0$  is the average implementation rate of the transition  $t$ , and it indicates the average implementation times per unit time. The inverse of the average implementation rate of the transition  $t_i$  is called average implementation delay or average service time.

There are three performance equivalence formulas of the series, parallel, and selection structured stochastic Petri net model, and the relevant conclusions are as follows:

**Theorem 1.** *The system B consists of n series transitions. Assume that the delay times of the n series transitions are the independent random variables, and they were subject to exponential distribution function with parameters  $\lambda_1, \lambda_2 \dots \lambda_n$ . So, the total equivalent delay time is as follows:*

$$\frac{1}{\lambda} = \sum_{i=1}^n \frac{1}{\lambda_i}. \quad (7)$$

**Theorem 2.** *The system B consists of n parallel transitions. Assume that the delay times of the n parallel transitions are the independent random variables, and they were subject to exponential distribution function with parameters  $\lambda_1, \lambda_2 \dots \lambda_n$ . So, the total equivalent delay time is as follows:*

$$\frac{1}{\lambda} = \sum_{i=1}^n \frac{1}{\lambda_i} - \sum_{i=1}^{n-1} \sum_{j=i+1}^{n-1} \sum_{k=j+1}^n \frac{1}{\lambda_i + \lambda_j + \lambda_k} + \dots \quad (8)$$

**Theorem 3.** *The system B consists of n selection transitions. Assume that the delay times of the n selection transitions are the independent random variables, and they were subject to exponential distribution function with parameters  $\lambda_1, \lambda_2 \dots \lambda_n$ . Assume that the implementation probability of the transition  $t_i$  is  $\alpha_i$ . So, the average delay time of the n selection transitions and the total equivalent delay time are as follows:*

$$\frac{1}{\lambda} = \sum_{i=1}^n \frac{\alpha_i}{\lambda_i}. \quad (9)$$

**3.2. Queuing Theory.** Queuing theory, also known as stochastic service system theory, is a mathematical discipline that studies crowding phenomenon. It solves the optimal system design and optimal control through the probabilistic characteristics of a variety of services in queuing system. Queuing theory is an important branch of operational research and an important branch of applied probability. The research has a strong practical background, which originated

by the Danish telecommunication engineer A. K. Erlang's research in the 20th century. Now, it is a mature theory.

Queueing phenomenon consists of two aspects: one obtains services, and the other tries to give services. The first one is called customer, which requires services. The servers give others services. The customers and servers constitute a queuing system.

**Definition 2.** The queuing system consists of three parts: the input process, queuing rules, and servers. These three parts of the queuing system are an organic whole, which describes the queuing system from different aspects.

- (1) The input process describes how the customer arrives at the queuing system. It is usually described by the probability distribution of the customer arrival time. The common distribution is as follows: fixed-length input, negative exponential input (Poisson stream), geometric input, Erlang input, and so on.
- (2) The queuing rules of the queuing system are the order of the customers accepting services. And it can be divided into first come first serve (FCFS), last come first serve (LCFS), and so on.
- (3) The structure of the server means the number of the servers and the service time of a customer. The common customer service time distributions are as follows: fixed-length distribution, the negative exponential distribution, and geometric distribution.

The queuing system model is as follows (see Figure 1). The mathematical representation of the queuing system is given in Theorem 4.

**Theorem 4.** *A queuing system can be expressed with a three-letter symbol A/B/C. A represents the customer arrival time distribution, B shows the distribution of the service time, and C represents the number of the servers. A and B will be M, D, and G. M represents a negative exponential distribution, D represents the distribution of fixed length, and G represents the general distribution.*

The key performance indicators of a queuing system are as follows:

- (1) The average queue length  $L$ , that is, the number of customers of the queuing system and its mathematical expectation formula, is as follows:

$$E[L] = \frac{\rho}{1 - \rho} = \frac{s}{\mu - s}. \quad (10)$$

- (2) The average staying time of the customer  $W$ , that is, the processing time of each customer in the system and its mathematical expectation formula, is as follows:

$$E[W] = \frac{1}{\mu - s}. \quad (11)$$

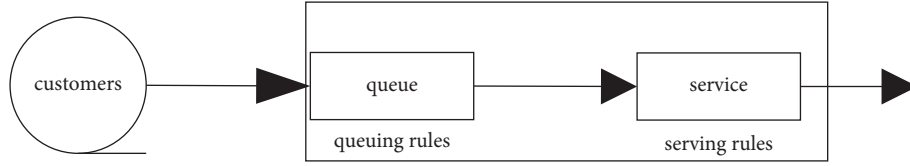


FIGURE 1: Queuing system model.

- (3) The average waiting time of the customer  $W_q$ , that is, the waiting time of each customer of the system and its mathematical expectation formula, is as follows:

$$E[W_q] = \frac{\rho}{\mu(1-\rho)}. \quad (12)$$

## 4. The STEP Event Stream Processing System

### 4.1. The Architecture of STEP Event Stream Processing System.

The architecture of spatial and temporal event stream processing system based on STEP in Internet of vehicles is shown in Figure 2. Firstly, the user writes the event query expressions by STEP and sends these expressions to the event stream processing system. Secondly, the event stream processing system receives the event query expressions and converts the expressions into event stream processing model. Then, the event stream processing system generates the corresponding event stream processing program according to the event stream processing model. The event stream processing program will begin processing the received event stream and producing the query results. Event stream processing program begins processing the received event stream and producing query results. Finally, the query results will be sent to the user.

To complete the event stream query, we need to convert the event query expression into the event stream processing model. There are four types of event stream processing models: tree model, directed acyclic graph model, automata model, and Petri net model. The Petri net model can be used to simulate the concurrency, asynchronous, distributed, nondeterministic, and parallelism systems. Therefore, we used Petri net to model the event processing system.

**4.2. The Core Algorithm of STEP Event Stream Processing System.** The main function of the STEP event stream processing system is to receive the event stream processing Petri net model and return the processing results. The core algorithm of the system is shown in Table 1.

In the algorithm, all input transitions of the Petri net model are divided into three categories: the beginning transition (TransitionStart), the middle transition (TransitionMid), and the end transition (TransitionEnd). The function of the TransitionStart is to detect the attribute values of the received event instance. The TransitionStart transmits the information through sharing the Place with the connected TransitionMid: the output Place of the TransitionStart shares the same variable with the input Place of the TransitionMid. The TransitionEnd is a special kind of transition: it has no successor.

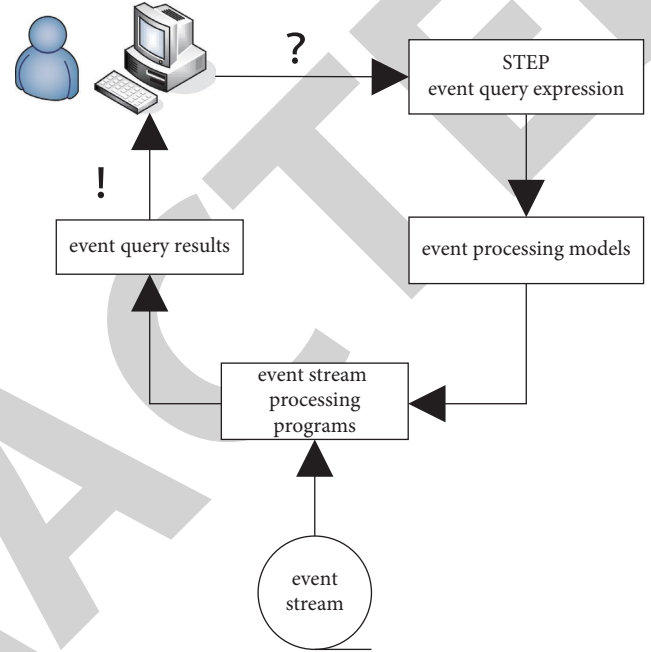


FIGURE 2: Architecture of the event stream processing system.

In the algorithm, the TransitionStart will be implemented firstly, and the other transitions will be implemented step by step until the TransitionEnd along the Petri net structure. Because the number of the transitions is limited from the TransitionStart to the TransitionMid, therefore, the algorithm is terminated for an event instance. Assume that there are  $n$  transitions in the Petri net model, and then the complexity of the algorithm is  $O(\log n)$ .

**4.3. The Queuing Theory Model of the STEP Event Stream Processing System.** The queuing theory model of the STEP event stream processing program based on Algorithm 1 is shown in Figure 3. It can clearly express the overall framework and processing process of Algorithm 1 in the queuing theory model.

We analyze the dynamic performance of the system by queuing theory. The dynamic performance of the queuing system mainly refers to the degree of change in the processing time of the service system when the speed of input data changes. Firstly, the event stream is transferred into the system memory, and these event instances will queue in there. Then, they will be processed by the event stream processing program one by one. The event stream in Internet of vehicles can be seen as Poisson stream. The service rule is first come first serve (FCFS), and the queuing system model of event stream processing system is an  $M/M/1$  queuing system.



TABLE 1: The core algorithm of the STEP event stream processing system.

Algorithm event stream processing algorithm based on Petri net model	
<b>Input:</b>	event stream, petri net model of the query expressions
<b>Output:</b>	query results
(1)	<b>while</b> TRUE <b>do</b>
(2)	Event $\leftarrow$ the pointer point the current event;
(3)	<b>foreach</b> Event.Attribute <b>do</b>
(4)	TransitionStart.InputToken $\leftarrow$ Event.Attribute;
(5)	<b>end</b>
(6)	<b>foreach</b> TransitionStart <b>do</b>
(7)	<b>if</b> TransitionStart.InputToken meet current guard <b>then</b>
(8)	OutputToken $\leftarrow$ InputToken;
(9)	<b>else</b>
(10)	End the query and jump to statement 28;
(11)	<b>end</b>
(12)	<b>end</b>
(13)	<b>foreach</b> TransitionMid <b>do</b>
(14)	<b>if</b> TransitionMid.InputToken meet guard <b>then</b>
(15)	OutputToken $\leftarrow$ OutputToken $\odot$ InputToken;
(16)	<b>else</b>
(17)	End the query and jump to statement 28;
(18)	<b>end</b>
(19)	<b>end</b>
(20)	<b>foreach</b> TransitionEnd <b>do</b>
(21)	<b>if</b> TransitionEnd.InputToken meet guard <b>then</b>
(22)	OutputToken $\leftarrow$ OutputToken $\odot$ InputToken;
(23)	<b>return</b> TransitionEnd.OutputToken;
(24)	<b>else</b>
(25)	End the query and jump to statement 28;
(26)	<b>end</b>
(27)	<b>end</b>
(28)	Delete the current event, the pointer point the next event;
(29)	<b>end</b>

There are two important parameters to analyze the dynamic performance of the queuing system: the parameter  $s$  of the Poisson event stream and the parameter  $\mu$  of the service time distribution of the event stream processing system.

Because the physical meaning of the parameter  $s$  of the Poisson event stream is the data transfer rate of networks, the parameter can be determined according to the network transmission environment. Next, we focus on the parameter  $\mu$  of the service time distribution.

The stochastic Petri net is a powerful tool to analyze the dynamic performance of the system. According to Definition 1, we can build the stochastic Petri net model of the event stream processing system. And we can analyze the dynamic performance of the event stream processing system by Theorem 4.

The example is shown as follows:

As shown in Figure 4, there are three independent random variables of the transition's delay time, and they were subject to parameters  $\lambda_1$ ,  $\lambda_2$ ,  $\lambda_3$  exponential distribution functions. The three random variables represent the processing time in each step of the computation in the data stream processing system.

Assume that the overall delay time of the stochastic Petri net obeys the exponential distribution function with parameter  $\lambda$ . According to the Theorem 4, we get

$$\frac{1}{\lambda} = \left( \frac{1}{\lambda_1} + \frac{1}{\lambda_2} - \frac{1}{\lambda_1 + \lambda_2} \right) + \frac{1}{\lambda_3}. \quad (13)$$

So, we can get the value of parameter  $\lambda$  in the end.

Assume that the processing rate  $\mu$  of event stream processing program is proportional to the speed  $\lambda$  of the stochastic Petri net model:  $\mu = k\lambda$  (where  $k > 0$  and the value of  $k$  is related to the processor speed of the operating environment). So, we can analyze the dynamic performance of an event stream processing program according to the speed parameter  $\mu$  of the event stream processing program and the transmission rate parameter  $s$  of the event stream.

## 5. Case Study

There are four different structured event stream processing program models, shown in Figure 5. And we analyze these examples by stochastic Petri net model and the queuing theory.

As shown in Figures 5(a)–5(d), there are four event stream processing programs: Example 1, Example 2, Example 3, and Example 4. In the examples, structure  $a$  represents a mixed calculation process of two AND operations. Structure  $b$  represents a mixed calculation process in which an OR operation is performed first, and then an AND operation is performed. Structure  $c$  represents a mixed calculation process in which an AND operation is performed first, and then an OR operation is performed. Structure  $d$  represents a mixed calculation process in which an OR operation is performed first, and then an AND operation is performed. In these examples, the transmission rate of the event stream is  $s_1$ ,  $s_2$ ,  $s_3$ , and  $s_4$ . Their event stream processing rate is  $\mu_1$ ,  $\mu_2$ ,  $\mu_3$ , and  $\mu_4$ . Because the service intensity of the queuing system  $\rho = s/\mu$ , and the necessary and sufficient conditions of the queuing system have the steady state distribution as  $\rho < 1$ , so assume that  $\rho_1 = s_1/\mu_1 < 1$ ,  $\rho_2 = s_2/\mu_2 < 1$ ,  $\rho_3 = s_3/\mu_3 < 1$ , and  $\rho_4 = s_4/\mu_4 < 1$ .

In order to compare these four structural examples' performance, we set the parameters as in Table 2. In the stream data processing structure of this paper, the former data processing generally takes less time, while the latter data processing takes more time. Therefore, according to this situation, we assume the parameter settings as shown in Table 2. According to Theorem 4, we can get the overall rate of these examples shown as the  $\lambda$  column of the table. Assume that the transition probability of each branch is equal in the selection structure.

According to the Theorem 4, the dynamic performance of the example 1 can be calculated as follows:

Assume that the event stream transmission rate is  $s_1$ , and the processing rate of the event processing program is  $\mu_1 = 33/23 k$ .

The queuing system's service intensity  $\rho_1 = s_1/\mu_1 = 23s_1/33k$ ;



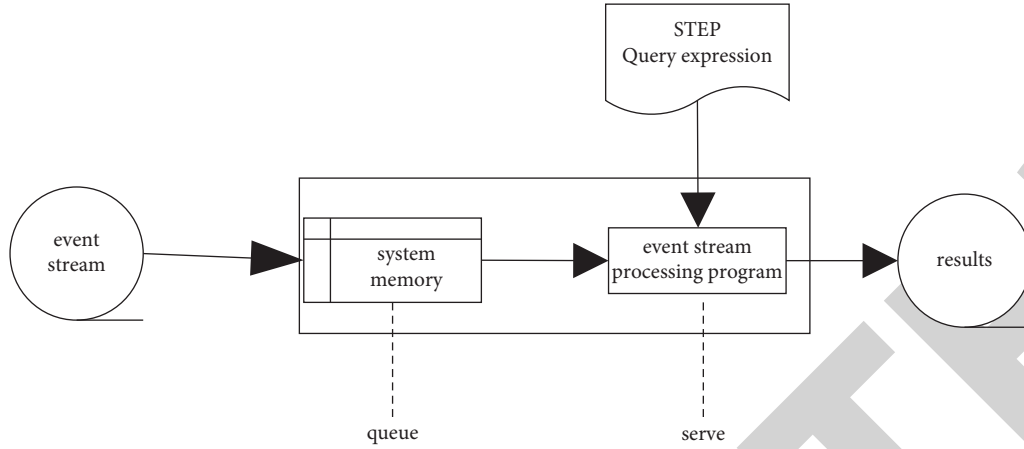


FIGURE 3: The queuing theory model of STEP event stream processing program.

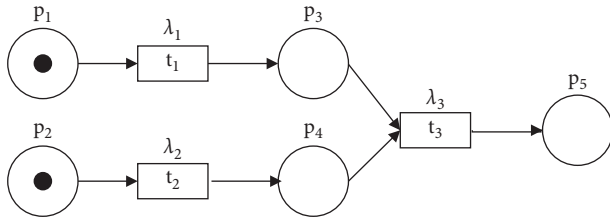


FIGURE 4: An example of event stream processing program's stochastic Petri net model.

The amount of the waiting events of the queuing system

$$E[L_1] = \rho_1/1 - \rho_1 = s_1/\mu_1 - s_1 = 23s_1/33k - 23s_1;$$

The residence time of each event of the queuing system

$$E[W_1] = 1/\mu_1 - s_1 = 23/33k - 23s_1;$$

The waiting time of each event of the queuing system

$$E[W_{q1}] = \rho_1/\mu_1(1 - \rho_1) = 529s_1/(33k - 23s_1)33k;$$

The proportion of waiting time of the queuing system

$$E[W_{q1}]/E[W_1] = 23s_1/33k;$$

Similarly, according to the Theorem 4, the dynamic performance of example 2 can be calculated as follows:

Assume that the event stream transmission rate is  $s_2$ , and the processing rate of the event processing program is  $\mu_2 = 16/5k$ ,

The queuing system's service intensity  $\rho_2 = s_2/\mu_2 = 5s_2/16k$ ;

The amount of the waiting events of the queuing system  $E[L_2] = \rho_2/1 - \rho_2 = s_2/\mu_2 - s_2 = 5s_2/16k - 5s_2$ ;

The residence time of each event of the queuing system  $E[W_2] = 1/\mu_2 - s_2 = 5/16k - 5s_2$ ;

The waiting time of each event of the queuing system  $E[W_{q2}] = \rho_2/\mu_2(1 - \rho_2) = 25s_2/(16k - 5s_2)16k$ ;

The proportion of waiting time of the queuing system  $E[W_{q2}]/E[W_2] = 5s_2/16k$ ;

Similarly, according to Theorem 4, the dynamic performance of example 3 can be calculated as follows:

Assume that the event stream transmission rate is  $s_3$ , and the processing rate of the event processing program is  $\mu_3 = 12/5k$ ,

The queuing system's service intensity  $\rho_3 = s_3/\mu_3 = 5s_3/12k$ ;

The amount of the waiting events of the queuing system  $E[L_3] = \rho_3/1 - \rho_3 = s_3/\mu_3 - s_3 = 5s_3/12k - 5s_3$ ;

The residence time of each event of the queuing system  $E[W_3] = 1/\mu_3 - s_3 = 5/12k - 5s_3$ ;

The waiting time of each event of the queuing system  $E[W_{q3}] = \rho_3/\mu_3(1 - \rho_3) = 25s_3/(12k - 5s_3)12k$ ;

The proportion of waiting time of the queuing system  $E[W_{q3}]/E[W_3] = 5s_3/12k$ ;

Similarly, according to the Theorem 4, the dynamic performance of example 4 can be calculated as follows:

Assume that the event stream transmission rate is  $s_4$ , and the processing rate of the event processing program is  $\mu_4 = 168/83k$ ,

The queuing system's service intensity  $\rho_4 = s_4/\mu_4 = 23s_4/33k$ ;

The amount of the waiting events of the queuing system  $E[L_4] = \rho_4/1 - \rho_4 = s_4/\mu_4 - s_4 = 83s_4/168k - 83s_4$ ;

The residence time of each event of the queuing system  $E[W_4] = 1/\mu_4 - s_4 = 83/168k - 83s_4$ ;

The waiting time of each event of the queuing system  $E[W_{q4}] = \rho_4/\mu_4(1 - \rho_4) = 6889s_4/(168k - 83s_4)168k$ ;

The proportion of waiting time of the queuing system  $E[W_{q4}]/E[W_4] = 83s_4/168k$ ;

In order to more clearly analyze the dynamic performance of the above examples, the discussion is now divided into two scenarios:

**5.1. Application Scene One.** The event stream's arrival rate  $s$  is changing, and the processing speed of the computer  $k$  is not changing. Assume that the constant  $k = 100$  and the dynamic performance of the system are as shown in Table 3.

The diagrams of the data in the Table 3 are shown in Figure 6. There are four subgraphs in the Figure 6. The diagram of the average amount of the waiting events in the queuing system of the four examples is Figure 6(a). The

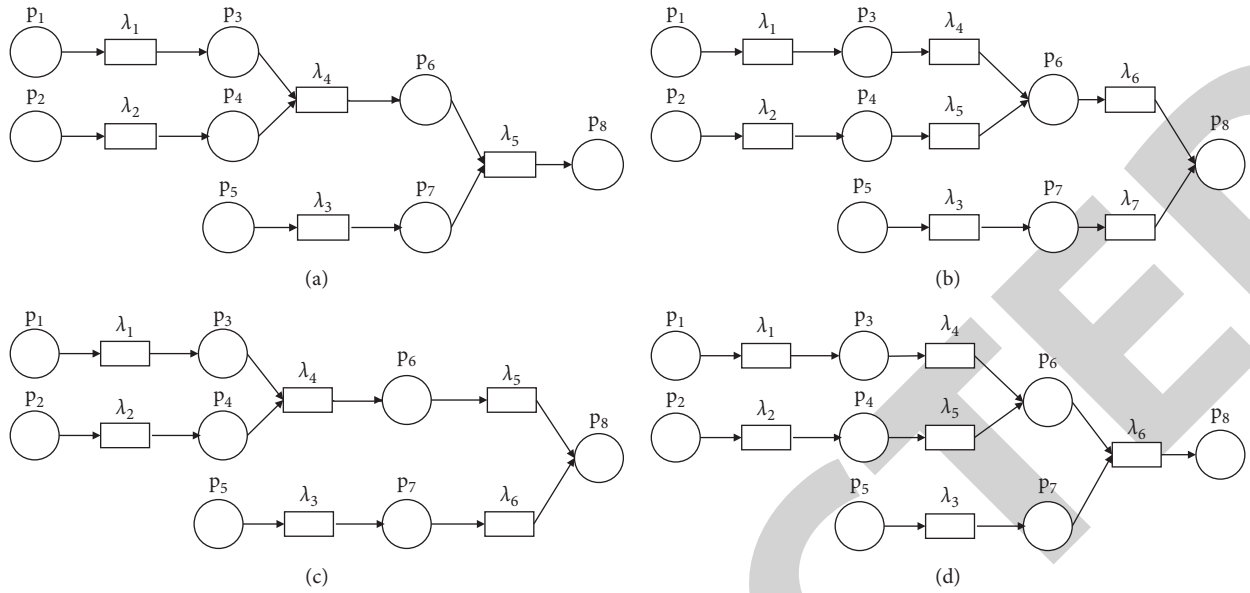


FIGURE 5: Four examples of event stream processing program's stochastic Petri net model.

TABLE 2: The parameters of four examples.

No.	$\lambda_1$	$\lambda_2$	$\lambda_3$	$\lambda_4$	$\lambda_5$	$\lambda_6$	$\lambda_7$	$\lambda$
Example 1	2	4	10	20	20	None	None	33/23
Example 2	2	4	10	20	20	20	20	16/5
Example 3	2	4	10	20	20	20	None	12/5
Example 4	2	4	10	20	20	20	None	168/83

diagram of the average residence time of the events in the queuing system of the four examples is Figure 6(b). The diagram of the average waiting time of the events in the queuing system of the four examples is Figure 6(c). And the diagram of the proportion of the waiting time of events in the queuing system of the four examples is Figure 6(d).

As can be seen in Figure 6 when the computer processing rate is not changing, so there is an increase in the event stream transmission rate, the average amount of the waiting events, the average residence time of the events, the average waiting time of the events, and the proportion of the waiting time of events, which is in the queuing system of the four examples. In addition, it can be seen from the figure that it takes the maximum processing event time in example 1 and the minimal processing event time in example 2.

5.2. *Application Scene Two.* The event stream's arrival rate  $s$  is not changing, and the processing speed of the computer  $k$  is changing. Assume that the constant  $s = 200$  and the dynamic performance of the system are as shown in Table 4.

The diagrams of the data in the Table 4 are shown in Figure 7. There are four subgraphs in Figure 7. The diagram of the average amount of the waiting events in the queuing system of the four examples is Figure 7(a). The diagram of the average residence time of the events in the queuing system of the four examples is Figure 7(b). The diagram of the average waiting time of the events in the queuing system

of the four examples is Figure 7(c). And the diagram of the proportion of the waiting time of events in the queuing system of the four examples is Figure 7(d).

As can be seen in Figure 7, when the event stream transmission rate is not changing, there is an increase in the computer processing rate, the average amount of the waiting events, the average residence time of the events, and the average waiting time of the events, and the proportion of the waiting time of events is reducing, which is in the queuing system of the four examples. In addition, it can be seen from the figure that it takes the maximum processing event time in example 1 and the minimal processing event time in example 2.

## 6. The Simulation and Results

In order to verify the derivation results of the dynamic performance of the previous section about the four examples, we build the four simulations of the examples by the queuing theory simulation tools Java Modelling Tools v0.9.1. For each example, the simulation program will run 100,000 times, and we get the average value of the results. In the simulation experiments, the boundary conditions are  $k = 100$  and  $s > 0$ .

6.1. *Simulation Experiment One.* The experiment corresponds to the application scenario one. When the event stream arrival rate is changing, and the computer processing

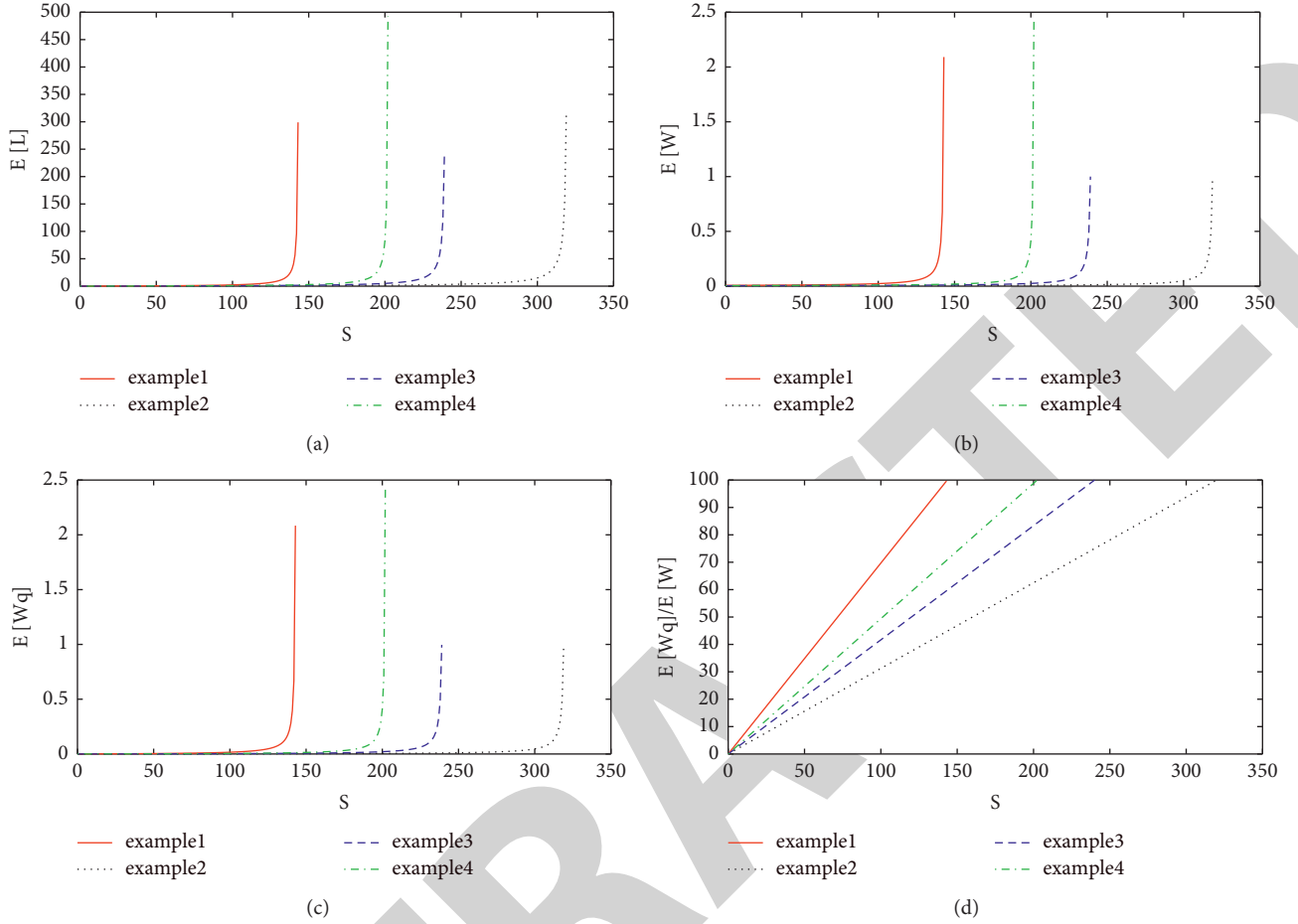


FIGURE 6: The diagram of dynamic performance of the examples in application scene one.

TABLE 3: The dynamic performance of the examples in application scene one.

No.	$\mu$	$s$	E (L)	E (W)	E (Wq)	$E[Wq]/E[W]$
Example 1	143.5	(0, 143.5)	$23s_1/3300 - 23s_1$	$23/3300 - 23s_1$	$529s_1/10890000 - 75900s_1$	$23s_1/3300$
Example 2	320	(0, 320)	$5s_2/1600 - 5s_2$	$5/1600 - 5s_2$	$25s_2/2560000 - 8000s_2$	$5s_2/1600$
Example 3	240	(0, 240)	$5s_3/1200 - 5s_3$	$5/1200 - 5s_3$	$25s_3/1440000 - 6000s_3$	$5s_3/1200$
Example 4	202.4	(0, 202.4)	$83s_4/16800 - 83s_4$	$83/16800 - 83s_4$	$6889s_4/282240000 - 1394400s_4$	$83s_4/16800$

TABLE 4: The dynamic performance of the examples in application scene two.

No.	$\mu$	$k$	E (L)	E (W)	E (Wq)	$E[Wq]/E[W]$
Example 1	(200, $+\infty$ )	(200, $+\infty$ )	$4600/33k_1 - 4600$	$23/33k_1 - 4600$	$105800/1089k_1^2 - 151800k_1$	$4600/33k_1$
Example 2	(200, $+\infty$ )	(62.5, $+\infty$ )	$125/2k_2 - 125$	$5/16k_2 - 1000$	$625/32k_2^2 - 2000k_2$	$125/2k_2$
Example 3	(200, $+\infty$ )	(83, $+\infty$ )	$250/3k_3 - 250$	$5/12k_3 - 1000$	$625/18k_3^2 - 1500k_3$	$250/3k_3$
Example 4	(200, $+\infty$ )	(98.8, $+\infty$ )	$2075/21k_4 - 2075$	$83/168k_4 - 16600$	$172225/3528k_4^2 - 348600k_4$	$2075/21k_4$

speed is not changing (here set  $k=100$ ), the simulation results are shown in Tables 5–8.

We plotted the simulation results in scatterplot and compared them with the diagram of theoretical reasoning results in Figure 8.

As can be seen from the above figure, the simulation results have agreed well with the previous theoretical results. In addition, as can be seen from the scatterplot of the experimental results, along with the event stream's arrival rate getting closed to the threshold, the event processing time has

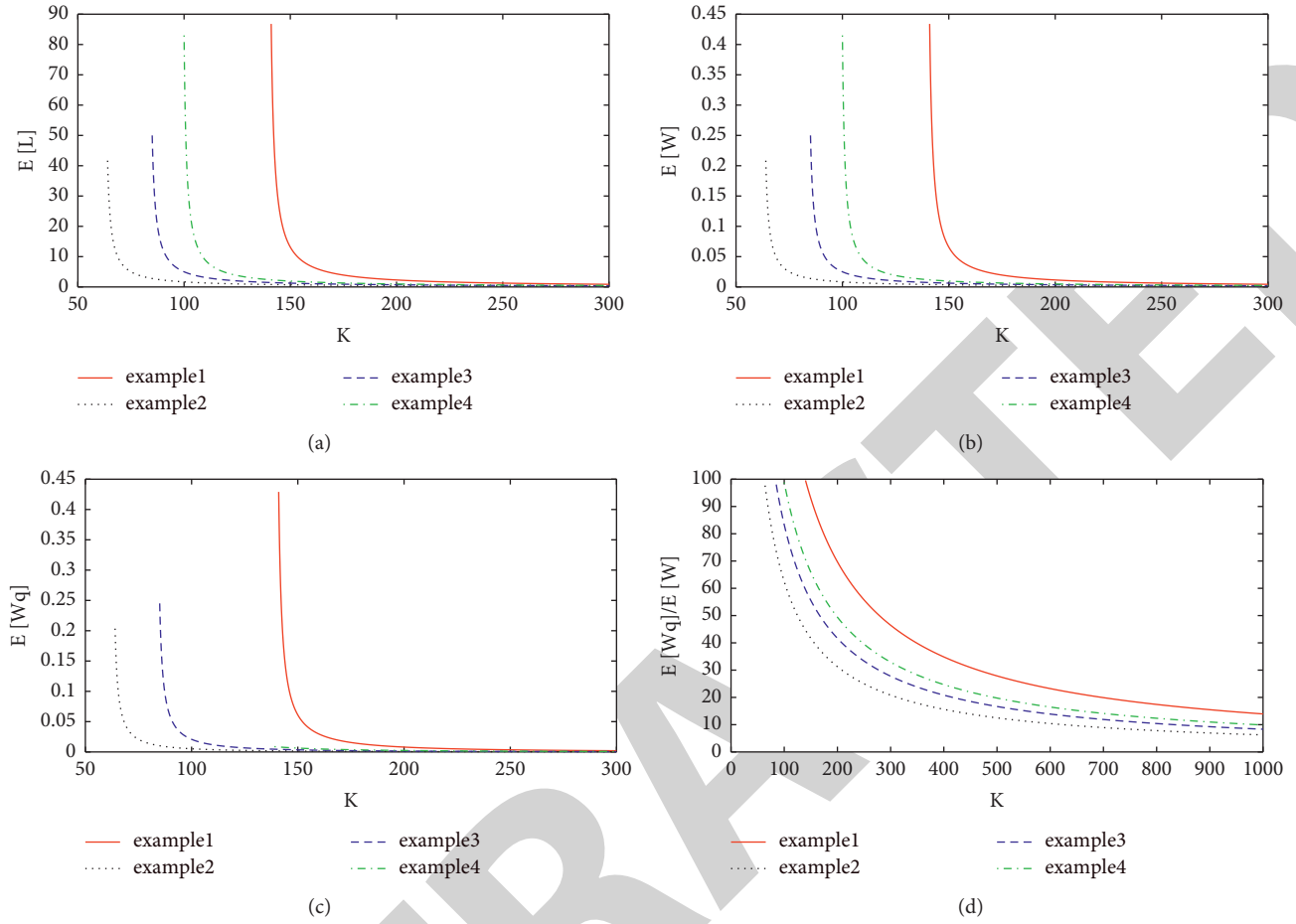


FIGURE 7: The diagram of dynamic performance of the examples in application scene two.

 TABLE 5: The results of example 1 of the simulation experiment one ( $\mu=143.5$ ).

No.	$S_1$	$E(L_1)$	$E(W_1)$	$E(W_{q1})$	$E[W_{q1}]/E[W_1]$	No.	$S_1$	$E(L_1)$	$E(W_1)$	$E(W_{q1})$	$E[W_{q1}]/E[W_1]$
01	14	0.106	0.0077	0.0008	9.72%	06	84	1.418	0.017	0.0098	57.76%
02	28	0.246	0.0087	0.0017	19.45%	07	98	2.169	0.022	0.015	68.18%
03	42	0.406	0.0096	0.0029	29.78%	08	112	3.532	0.032	0.025	78.13%
04	56	0.64	0.011	0.0044	40.09%	09	126	7.545	0.059	0.053	89.83%
05	70	0.955	0.014	0.0068	48.36%	10	135	15.621	0.114	0.107	93.86%

 TABLE 6: The results of example 2 of the simulation experiment one ( $\mu=320$ ).

No.	$S_2$	$E(L_2)$	$E(W_2)$	$E(W_{q2})$	$E[W_{q2}]/E[W_2]$	No.	$S_2$	$E(L_2)$	$E(W_2)$	$E(W_{q2})$	$E[W_{q2}]/E[W_2]$
01	32	0.112	0.0035	0.0003	9.94%	06	192	1.498	0.0078	0.0048	60.66%
02	64	0.249	0.0039	0.0008	20.03%	07	224	2.318	0.01	0.0073	72.9%
03	96	0.43	0.0044	0.0014	30.61%	08	256	3.999	0.016	0.012	75%
04	128	0.662	0.0052	0.0021	39.46%	09	288	9.404	0.031	0.028	90.32%
05	160	0.987	0.0062	0.0032	50.64%	10	310	30.149	0.099	0.096	96.97%

dramatically increased. But the ratio of the waiting time and the processing time has uniformly changed.

**6.2. Simulation Experiment Two.** The experiment corresponds to the application scenario one. When the event stream arrival rate is not changing (here set  $s=200$ ), and the

computer processing speed is changing, the simulation results are shown in Tables 9–12.

We plotted the simulation results in scatterplot and compared them with the diagram of theoretical reasoning results as shown in Figure 9.

As can be seen from the above figure, the simulation results have agreed well with the previous theoretical results.

TABLE 7: The results of example 3 of the simulation experiment one ( $\mu = 240$ ).

No.	$S_3$	$E(L_3)$	$E(W_3)$	$E(W_{q_3})$	$E[W_{q_3}]/E[W_3]$	No.	$S_3$	$E(L_3)$	$E(W_3)$	$E(W_{q_3})$	$E[W_{q_3}]/E[W_3]$
01	24	0.112	0.0046	0.0005	10.11%	06	144	1.494	0.01	0.0062	61.7%
02	48	0.252	0.0052	0.001	20%	07	168	2.312	0.014	0.0097	69.57%
03	72	0.429	0.006	0.0018	29.95%	08	192	3.959	0.021	0.017	80.95%
04	96	0.663	0.0069	0.0028	39.97%	09	216	8.649	0.041	0.036	87.8%
05	120	0.987	0.0083	0.0041	49.94%	10	230	20.773	0.097	0.093	95.88%

TABLE 8: The results of example 4 of the simulation experiment one ( $\mu = 202.4$ ).

No.	$S_4$	$E(L_4)$	$E(W_4)$	$E(W_{q_4})$	$E[W_{q_4}]/E[W_4]$	No.	$S_4$	$E(L_4)$	$E(W_4)$	$E(W_{q_4})$	$E[W_{q_4}]/E[W_4]$
01	20	0.11	0.0055	0.0005	9.89%	06	120	1.444	0.012	0.0072	59.75%
02	40	0.244	0.0061	0.0012	20.03%	07	140	2.255	0.016	0.011	68.75%
03	60	0.42	0.007	0.0021	30.1%	08	160	3.786	0.024	0.019	79.17%
04	80	0.652	0.0082	0.0032	39.19%	09	180	8.042	0.044	0.04	90.91%
05	100	0.978	0.0097	0.0048	49.02%	10	195	24.98	0.135	0.13	96.3%

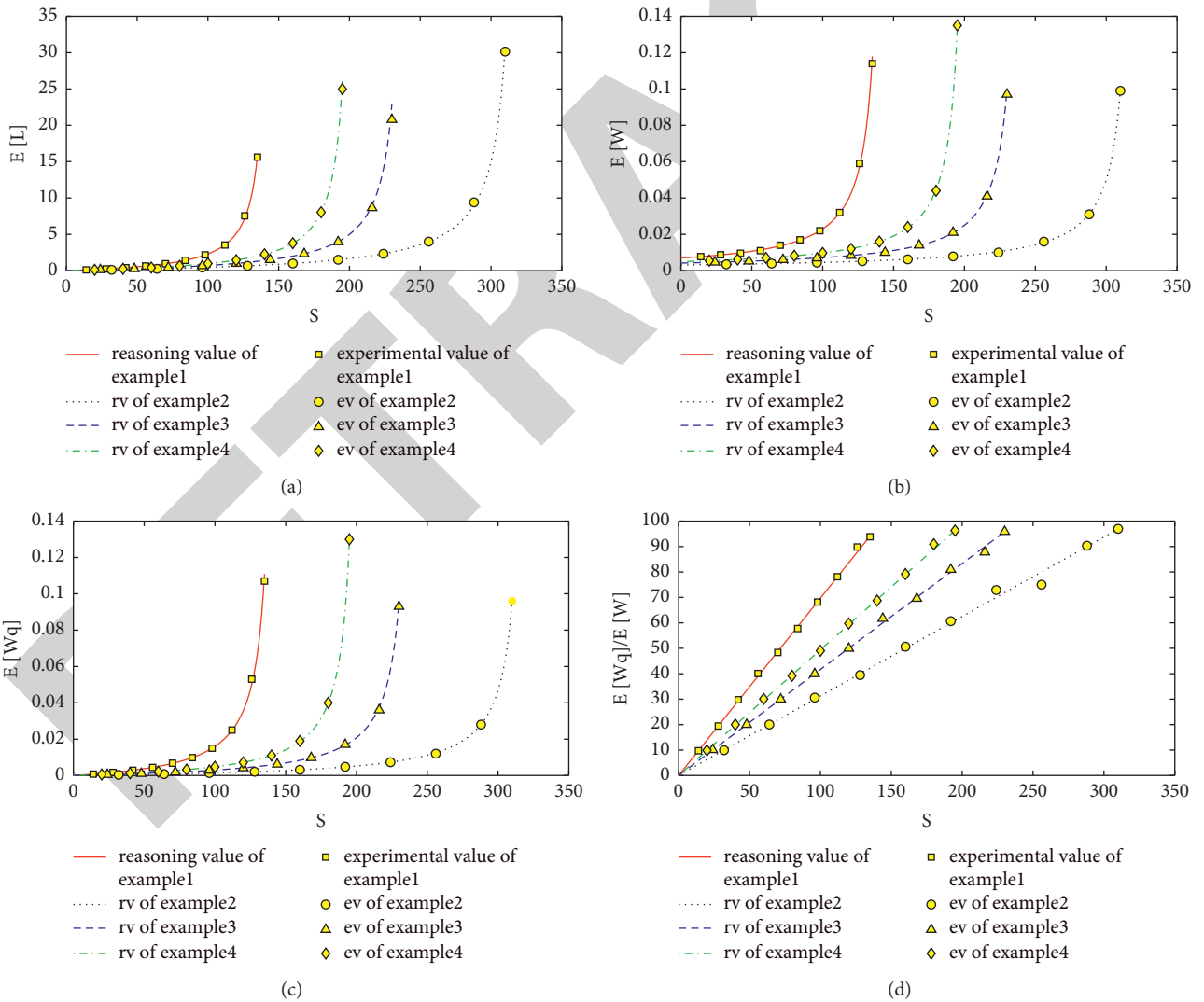


FIGURE 8: The comparison chart of dynamic performance of the examples in application scene one.

TABLE 9: The results of example 1 of the simulation experiment two.

No.	$\mu_1$	E ( $L_1$ )	E ( $W_1$ )	E ( $W_{q_1}$ )	E[ $W_{q_1}$ ]/E[ $W_1$ ]	No.	$\mu_1$	E ( $L_1$ )	E ( $W_1$ )	E ( $W_{q_1}$ )	E[ $W_{q_1}$ ]/E[ $W_1$ ]
01	215	12.737	0.062	0.059	95.16%	06	359	1.248	0.0063	0.0035	54.82%
02	244	4.402	0.022	0.018	81.82%	07	387	1.064	0.0053	0.0028	52.64%
03	273	2.768	0.014	0.01	71.43%	08	416	0.925	0.0046	0.0022	48.03%
04	301	1.968	0.0099	0.0066	66.13%	09	445	0.826	0.0041	0.0019	45.37%
05	330	1.527	0.0076	0.0047	61.26%	10	473	0.722	0.0036	0.0015	42.27%

TABLE 10: The results of example 2 of the simulation experiment two.

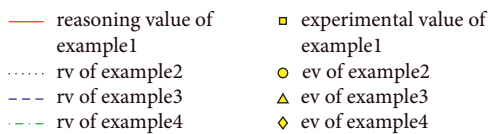
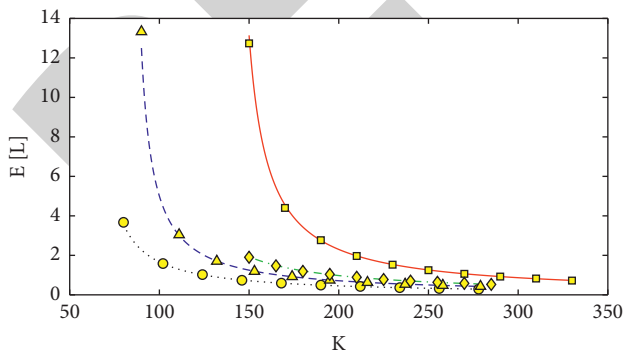
No.	$\mu_2$	E ( $L_2$ )	E ( $W_2$ )	E ( $W_{q_2}$ )	E[ $W_{q_2}$ ]/E[ $W_2$ ]	No.	$\mu_2$	E ( $L_2$ )	E ( $W_2$ )	E ( $W_{q_2}$ )	E[ $W_{q_2}$ ]/E[ $W_2$ ]
01	256	3.672	0.018	0.014	77.78%	06	608	0.493	0.0025	0.0008	32.54%
02	326	1.585	0.0079	0.0048	61.19%	07	678	0.425	0.0021	0.0006	29.53%
03	397	1.025	0.0052	0.0026	50.29%	08	749	0.364	0.0018	0.0005	26.14%
04	467	0.738	0.0037	0.0016	42.23%	09	819	0.321	0.0016	0.0004	24.66%
05	538	0.592	0.0029	0.0011	37.76%	10	890	0.288	0.0014	0.0003	22.71%

TABLE 11: The results of example 3 of the simulation experiment two.

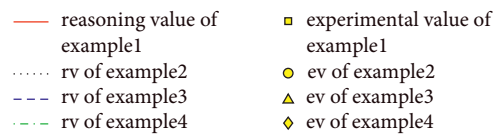
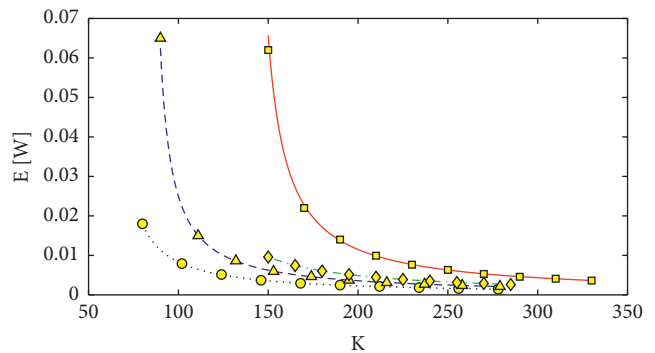
No.	$\mu_3$	E ( $L_3$ )	E ( $W_3$ )	E ( $W_{q_3}$ )	E[ $W_{q_3}$ ]/E[ $W_3$ ]	No.	$\mu_3$	E ( $L_3$ )	E ( $W_3$ )	E ( $W_{q_3}$ )	E[ $W_{q_3}$ ]/E[ $W_3$ ]
01	216	13.328	0.065	0.059	90.77%	06	468	0.748	0.0037	0.0016	42.9%
02	266	3.044	0.015	0.012	80%	07	518	0.628	0.0031	0.0012	39.42%
03	317	1.703	0.0086	0.0053	61.88%	08	569	0.536	0.0027	0.001	35.59%
04	367	1.189	0.0059	0.0033	55.05%	09	619	0.48	0.0024	0.0008	31.88%
05	418	0.924	0.0046	0.0022	47.82%	10	670	0.429	0.0021	0.0006	30.19%

TABLE 12: The results of example 4 of the simulation experiment two.

No.	$\mu_4$	E ( $L_4$ )	E ( $W_4$ )	E ( $W_{q_4}$ )	E[ $W_{q_4}$ ]/E[ $W_4$ ]	No.	$\mu_4$	E ( $L_4$ )	E ( $W_4$ )	E ( $W_{q_4}$ )	E[ $W_{q_4}$ ]/E[ $W_4$ ]
01	304	1.903	0.0096	0.0062	65.14%	06	455	0.777	0.0039	0.0017	44.1%
02	334	1.469	0.0074	0.0045	60.83%	07	486	0.694	0.0035	0.0014	40.97%
03	364	1.187	0.006	0.0033	55.24%	08	516	0.635	0.0031	0.0012	39.3%
04	395	1.037	0.0051	0.0026	50.58%	09	547	0.583	0.0029	0.0011	36.77%
05	425	0.891	0.0045	0.0021	46.98%	10	577	0.528	0.0026	0.0009	34.039%



(a)



(b)

FIGURE 9: Continued.



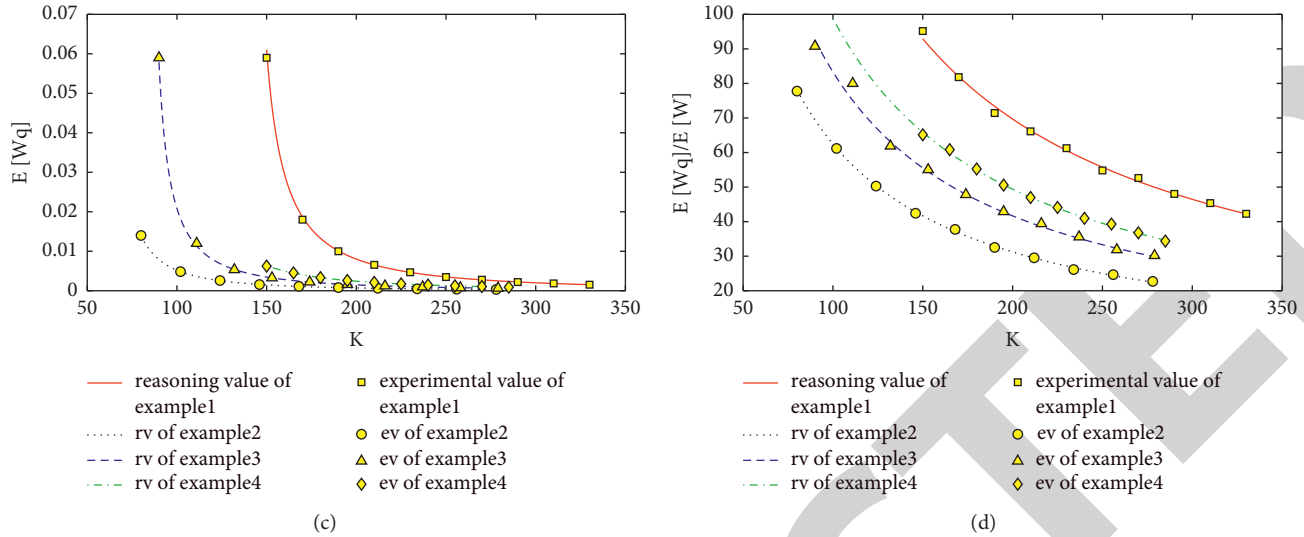


FIGURE 9: The comparison chart of dynamic performance of the examples in application scene two.

In addition, as can be seen from the scatterplot of the experimental results, along with the computing speed increasing, the event processing time has quickly reduced. But the ratio of the waiting time and the processing time has uniformly changed.

## 7. Conclusions

In this paper, we proposed the architecture and the core algorithm of STEP event stream processing system at first. Secondly, we analyzed the dynamic performance of the event stream processing system with stochastic Petri net and queuing theory. These dynamic performances include the queue length of the event stream, the processing time of each event instance, the waiting time of each event instance, and the proportion of the event waiting time. Finally, we gave four structured examples and gave their theoretical dynamic performance. We also built their simulation program to verify the correctness of the reasoning results, which shows the effectiveness of the relevant methods.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported by the 2021 Basic Science Research Program of Nantong, Jiangsu, China (No. JC2021127) and the Science Foundation of Jiangsu Shipping College, China (No. HYKY/2021B07).

## References

- [1] X. Xu, H. Li, W. Xu, Z. Liu, L. Yao, and F. Dai, "Artificial intelligence for edge service optimization in internet of vehicles: a survey," *Tsinghua Science and Technology*, vol. 27, no. 2, pp. 270–287, 2021.
- [2] P. Szurgott and P. Bernacki, "Modelling of steel-concrete bridges subjected to a moving high-speed train," *International Journal of Simulation Modelling*, vol. 19, no. 1, pp. 29–40, 2020.
- [3] Y. Yang, "A vehicle recognition algorithm based on deep convolution neural network," *Traitement du Signal*, vol. 37, no. 4, pp. 647–653, 2020.
- [4] G. W. Hua and Y. D. Xu, "Optimal deployment of charging stations and movable charging vehicles for electric vehicles," *Journal of System and Management Sciences*, vol. 9, no. 1, pp. 105–116, 2019.
- [5] D. N. Nguyen, T. A. Nguyen, T. B. Hoang, and N. D. Dang, "Establishing the method to predict the limited roll angle of the vehicle based on the basic dimensions," *Mathematical Modelling of Engineering Problems*, vol. 8, no. 5, pp. 775–779, 2021.
- [6] V. Praneeth, K. R. Kumar, and N. Karyemsetty, "Security: intrusion prevention system using deep learning on the internet of vehicles," *International Journal of Safety and Security Engineering*, vol. 11, no. 3, pp. 231–237, 2021.
- [7] M. Naeem, T. Jamal, J. Diaz-Martinez et al., "Trends and future perspective challenges in big data," *Advances in Intelligent Data Analysis and Applications*, pp. 309–325, 2022, Springer, Singapore.
- [8] A. Toshniwal, S. Taneja, A. Shukla et al., "Storm@ twitter," in *Proceedings of the 2014 ACM SIGMOD international conference on Management of data*, pp. 147–156, June 2014.
- [9] L. Neumeyer, B. Robbins, A. Nair, and A. Kesari, "S4: distributed stream computing platform," in *Proceedings of the IEEE International Conference on Data Mining Workshops*, pp. 170–177, Sydney, NSW, Australia, December 2010.
- [10] D. Borthakur, J. Gray, J. S. Sarma et al., "Apache hadoop goes realtime at facebook," in *Proceedings of the 2011 ACM SIGMOD International Conference on Management of data*, pp. 1071–1080, June 2011.

## Retraction

# Retracted: Design of Sponsored Search Auction Mechanism for Federated Learning Advertising Platform

### Computational Intelligence and Neuroscience

Received 1 August 2023; Accepted 1 August 2023; Published 2 August 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] H. Jiang, T. Cui, and K. Yang, "Design of Sponsored Search Auction Mechanism for Federated Learning Advertising Platform," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 5787491, 15 pages, 2022.

## Research Article

# Design of Sponsored Search Auction Mechanism for Federated Learning Advertising Platform

Hong Jiang <sup>1</sup>, Tianxu Cui <sup>2</sup> and Kaiwen Yang<sup>2</sup>

<sup>1</sup>School of Management, Shenzhen Polytechnic, Shenzhen 518055, China

<sup>2</sup>School of Information, Beijing Wuzi University, Beijing 101149, China

Correspondence should be addressed to Hong Jiang; [jianghong@szpt.edu.cn](mailto:jianghong@szpt.edu.cn) and Tianxu Cui; [cuitianxubwu@163.com](mailto:cuitianxubwu@163.com)

Received 19 January 2022; Revised 28 January 2022; Accepted 14 February 2022; Published 7 April 2022

Academic Editor: Daqing Gong

Copyright © 2022 Hong Jiang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Federated learning has demonstrated strong capabilities in terms of addressing concerns related to data islands and privacy protection. However, in real application scenarios, participants in federated learning have difficulty matching. For example, two companies distributed in different regions do not know that the other party also needs federated learning in the case of information asymmetry. Therefore, it is difficult to build alliances. To enable suppliers and consumers to find one or more federated learning objects that are relatively satisfactory in a short time, this paper considers the idea of establishing a federated learning advertising platform, where data transactions need to consider privacy protection. A sponsored search auction mechanism design method is introduced to solve the problem of ranking the presentation order of participant advertisements. Due to the potential malicious bidding problem, which occurs when using the classic sponsored search auction mechanism under the federated learning scenario, this paper proposes a novel federated sponsored search auction mechanism based on the Myerson theorem, improving upon the ranking index used in the classic sponsored search auction mechanism. A large number of experimental results on a simulation data set show that our proposed method can fairly select and rank the data providers participating in the bidding. Compared with other benchmark mechanisms, the malicious bidding rate is significantly decreased. In the long run, the proposed mechanism can encourage more data providers to participate in the federated learning platform, thus continuously promoting the establishment of a federated learning ecosystem.

## 1. Introduction

In recent years, with the application of artificial intelligence in all walks of life [1], data privacy protection has become increasingly valued by individuals and organizations. On the one hand, people are unwilling to share data that contain personal information any more, due to the potential risk of privacy leakage. On the other hand, regulatory agencies in various regions and countries have also issued a series of policies and regulations to protect the security of private data. For example, the General Data Protection Regulation (GDPR) [2], officially implemented by the European Union in May 2018, puts forward the most stringent requirements for data privacy and security. This is currently the most comprehensive and widely used privacy protection law in the world. In China, the

“Implementation Regulations of the Consumer Protection Law of the People’s Republic of China”, promulgated in August 2016, applies to most companies that deal with consumer data, forcing companies to take responsibility for their obligations to protect consumer personal information.

Federated Learning (FL) has demonstrated strong capabilities in addressing data islands and privacy protection. Within the past few years, federated learning has developed rapidly, in terms of algorithms, frameworks, platforms, and applications, since Google first proposed the concept in 2016. To incentivize high-quality data providers to contribute their data to the federation, researchers have designed different incentive mechanisms. However, in the real environment, data are stored at the edge nodes of various institutions.

For a certain federated learning task, some uncertain information exists in the following aspects:

- (1) Which data providers have the willingness to contribute data? What is the sample size that they can provide? What are the features? How much will it cost?
- (2) Which data users have a demand for data? What types of data do the data users need? What is the quantity demand? How much will they pay for it?

To enable suppliers and consumers to find one or more FL objects that are relatively satisfactory in the shortest time, it is necessary to provide a Federated Learning advertising platform. The data provider, on the supply side, can publish a brief introduction about their data on the advertising platform. Data users, on the demand side, can also find their intended customers on the platform. After matching, the data provider and the data user will train a machine learning model together using federated learning. As shown in Figure 1, the federated learning advertising platform provides advertising slots and charges advertising fees. Data providers post advertisements on the platform and pay the fees. Data users then browse the advertisements, and send out invitations to intended customers for federated learning.

Generally speaking, some data providers want their ads to appear in the most attractive slot: the first place in the result page. However, the slots for displaying advertisements are limited. Therefore, for each search, the FL advertising platform faces the problem of matching advertisements with slots. In addition, the FL advertising platform also needs to determine the price that each data provider needs to pay. Data providers naturally prefer eye-catching slots. Therefore, to allocate a slot and determine the corresponding bid, the FL advertising platform needs a ranking system. There are an increasing number of data providers who want to post ads on the platform, but the number of available ad slots remains the same. To solve this problem, we can use an auction mechanism. These auctions are called Sponsored Search Auctions [3]. In a typical sponsored search auction, the data provider is invited to bid on the keywords, that is, for each click on the advertisement, the advertiser expresses the maximum amount they are willing to pay. This is usually referred to as cost-per-click (CPC). According to the bid submitted by the advertisers for a specific keyword, the search engine selects a set of advertisements and determines their presentation order. The actual price of the search engine also depends on the bid submitted by the customers. However, the sponsored search auction mechanism (GFP, GSP, VCG) mentioned above is not directly applicable to the scenario of federated learning as, if the bid is the only factor considered in the allocations of data providers, it may lead to malicious bidding and the emergence of monopoly issues. Many data providers which have noisy data in their data set will make higher bids to occupy the top slots on the advertising platform. Obviously, this is irresponsible to both the suppliers and consumers. In the long run, the advertising platform will suffer serious losses.

To resolve the above problems, we propose a new mechanism, called Federated Sponsored Search Auction Mechanism (FSSA), which is specific to the scenario of federated learning. Based on the bids, this paper also adds an

index to reflect the average contribution of data providers in their previous federated learning scenarios, and assign certain weights to these two indices. Thus, a new ranking index is formally defined, ranking scores, which is more in line with the scenario of federated learning. The core idea is to consider both the bid and contribution to rank the data providers, such that some data providers whose data quality is poor but have strong economic strength can gain higher ad slots in the advertising platform; in the same way, some participants who lack money but who have better data quality can also rank highly on the advertising platform. Most importantly, this mechanism curbs the monopoly problem, to a certain extent, and secures the platform's long-term prosperity. Extensive experimental results on a simulation data set show that the FSSA can filter and rank the data providers participating in the bidding. Compared with the classic sponsored search auction mechanism, the malicious bidding rate can be significantly decreased, effectively avoiding the problem of inconsistency between the bid and the actual contribution ability.

Our contributions can be summarized as follows:

- (i) In response to the difficulty of matching data providers and data users in federated learning, we first propose the idea of establishing an FL advertising platform. This platform can help both the suppliers and consumers to find a relatively satisfactory federated learning object in the shortest time. Furthermore, a new ranking mechanism is proposed to solve the problem of ranking advertisement slots.
- (ii) For the classic sponsored search auction mechanism, only bids are considered for the ranking, thus leading to malicious bids. We propose a new ranking index—ranking scores—which considers both the bids and historical average contributions of federated learning participants. To a certain extent, the problem of malicious bids can be avoided.
- (iii) We conduct extensive experiments on different simulation data sets to verify the effectiveness of our proposed method. The experimental results show that our method can effectively rank data providers. Compared with the classic sponsored search auction mechanism, the rate of malicious bidding is significantly decreased.

The remainder of this paper is arranged as follows: We summarize the related literature in Section 2. The method and results are detailed in Section 3. In Section 4, we conclude the paper and discuss some prospects for future work. To the best of our knowledge, this is the first time that the theory of sponsored search auctions has been applied to a federated learning scenario. The factors which we consider are not only the bids of data providers but also a comprehensive survey of their contributions. A novel mechanism is proposed to ensure that the data providers of the FL advertising platform are treated fairly, as well as to encourage more data providers to participate in the federated learning



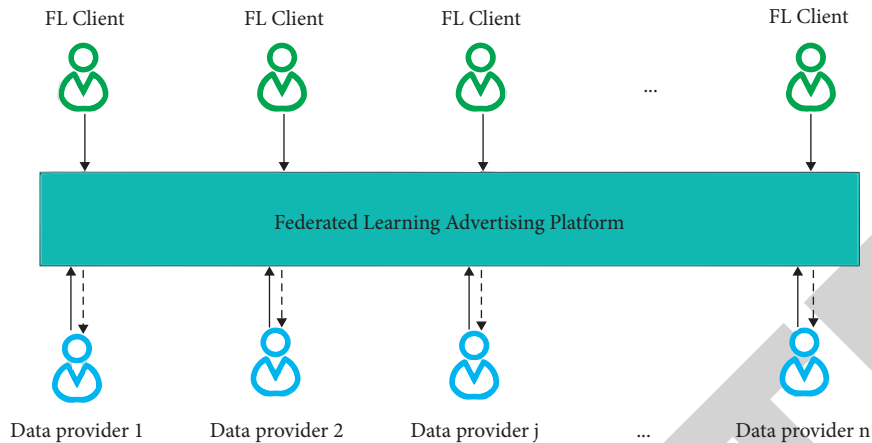


FIGURE 1: Schematic diagram of the Federated Learning Advertising Platform.

platform, in order to promote the establishment of the federated learning ecology.

## 2. Literature Review

We review the related literature from the following three aspects: Federated learning, mechanism design, and sponsored search auctions.

**2.1. Federated Learning.** Federated learning is a distributed machine learning framework [4, 5], which aims to solve the problems of user privacy and data islands which occur in the process of machine learning. Without data transmission, this method can train a machine learning model through using data from various devices. Since Google first proposed federated learning in 2016 [6], it has been favored by many technology companies, and many specialized research teams have been established. Not only have they achieved a lot of research results in frameworks and algorithms but also have actively explored and advanced the business landing applications of federated learning. Yang et al. [7] have expanded the concept and application of federated learning. According to the distribution features of data islands, federated learning can be divided into three types: horizontal federated learning, vertical federated learning, and federated transfer learning.

Recent research into federated learning has mainly focused on privacy protection and incentive mechanisms. To prevent the leakage of private data during gradient and parameter sharing in federated learning, researchers have applied traditional privacy protection techniques to federated learning, mainly divided into the following aspects: Secure multi-party computing [8–10], homomorphic encryption [11, 12], and differential privacy [13–15]. Some studies have begun to use the immutability of blockchain technology [16, 17] for privacy protection in federated learning. In terms of incentive mechanism design, Cong et al. [18] have established a research framework for the design and reasoning of federated learning incentive mechanisms, and proposed a precise definition of the FML incentive mechanism design problem. They divided this big problem into demand and supply-side problems for research and design. Zeng et al. [19] conducted a comprehensive research review on the design of

federated learning incentive mechanisms. They showed that the Federated Shapley value retains the desirable properties of the canonical Shapley value; it can be calculated without incurring additional communication costs, and it can also capture the influence of the participation order on data value. Different from the above works, Wei et al. [20] have proposed the concept of contribution index—a new metric based on Shapley value—which is suitable for evaluating the contribution of each data provider in the joint model of joint learning training. To solve the problem related to the large amount of calculations needed for the contribution index, they proposed two gradient-based methods. In addition, many scholars have introduced auction theory into the design process of the reward mechanism of federated learning. Kim et al. [21] have studied the incentive mechanism and privacy protection of federated learning from the perspective of mechanism design. Jiao et al. [22] have proposed an auction-based market model to encourage data owners to participate in joint learning, and designed two auction mechanisms for the federated learning platform to maximize the social welfare of the federated learning service market. In this paper, when defining the ranking scores of the FL advertising platform, we use a federated learning participant contribution index, which was defined and used by Wei et al. [20]. The contribution of each participant in federated learning can be measured fairly, to a certain extent, which, to some degree, represents the quality of the data held by the participant.

**2.2. Mechanism Design.** Hurwicz [23] first proposed the concept of mechanism design in the 1960s. He defined mechanism design as systems that can communicate with each other. It is the process of assigning results to participants, based on pre-made rules and information received in each round. In 1961, Vickrey published a paper on the second-price sealed-bid auction (i.e., a Vickrey Auction) [24], which was a milestone in the field of mechanism design. Three papers on incomplete information game theory, published by Harsanyi in the 1960s [25, 27], laid a solid foundation for mechanism design. Hurwicz introduced the concept of incentive compatibility into the field of mechanism design in 1972, thus opening up a period of rapid development in mechanism design [28]. Soon after, Clarke

[29] and Groves [30] extended the Vickrey mechanism to a more general quasilinear environment, and developed a more general incentive compatibility. In the 1970s, thanks to the joint efforts of a group of outstanding scholars, such as Gibbard [31] and Maskin [32], mechanism design made great progress in the revelation principle and implementation theory. As time progressed, mechanism design has been applied to many different disciplines, such as auctions, design of markets and trading institutions [33, 34], social choice theory [35], computer science [3], and so on.

**2.3. Sponsored Search Auctions.** The auction is the core content of mechanism design research, and sponsored search auctions have been a hotspot in the field of computer science over the past decade [36]. For traditional revenue optimization, Myerson [37] solved the single-term problem in the Bayesian–Nash equilibrium environment. Krysta et al. [38] have studied the multi-unit optimal auction. Since the first sponsored search auction was initiated in 1997, a series of auction mechanisms have been proposed, such as Overture’s Generalized First Price (GFP) auction [39] and Google’s Generalized Second Price (GSP), which has caused sponsored search auctions to become an important source of income for online platforms. Due to the inauthenticity of GSP, Garg et al. [40], based on the previous work, modeled the sponsored search auction on the Internet as a mechanism design problem, and designed a novel optimal auction mechanism (OPT), which can maximize the expected benefits of search engines while achieving Bayesian incentive compatibility and individual rationality. Lahaie et al. [41] have proposed another idea of compressing GSP parameters to increase income. Ostrovsky et al. [42] applied these works and studied the impact on Yahoo using auctions with the best reserve price. Zhang et al. [43] proposed an online reverse auction scheme for cloud computing service allocation based on Vickrey–Clarke–Groves (VCG) mechanism and online algorithm (OA), which can help cloud users and providers to build workflow applications in the cloud computing environment. This analytical approach has important implications for measuring the performance of the proposed algorithm without assuming the distribution of cloud provider bids.

As for the trade-off between different goals, some related studies can be found in the literature. Sundararajan et al. [44] have considered a convex combination of income and welfare to improve the forecast. Li et al. [45] constructed an integrated system with a mixed arrangement of advertising and organic items, and determined the best trade-off between instant income and user experience. In addition, for different settings and actual requirements, they extended the proposed optimal truthful allocation mechanism to meet these realistic conditions. With the help of real data, they verified the advantages of the proposed mechanism over the classic Myerson optimal advertising mechanism. Lian et al. [46] have optimized the advertising pruning of sponsored search based on reinforcement learning. This is the first time that reinforcement learning technology has been used to address this problem. More importantly, it has been

successfully implemented in Baidu’s sponsored search system, and online long-term A/B tests have shown a significant increment in revenue.

### 3. Method and Results

#### 3.1. Preliminary Knowledge and definition

##### 3.1.1. Typical Service Process of FL Advertising Platform

*Step 1.* As shown in Figure 2, the federated learning client enters keywords to search for a certain learning task on the advertising platform.

*Step 2.* After receiving the keywords, the federated learning advertising platform ranks the  $n$  bids  $b^{(j)}$  associated with the keywords of data providers using the ranking allocation rules, which are set in advance. After ranking,  $m$  data providers are selected, who are then ranked according to their ranking scores  $r^{(j)}$ .

*Step 3.* The customer browses the data set information provided by data providers on the advertising platform, selecting one or more data providers that they are satisfied with. Then, the client sends them an invitation for federated learning, and matches them according to their intentions to form a federated learning alliance.

*Step 4.* The paired data providers and customers conduct federated learning as an alliance, and the benefits are allocated according to the profit allocation rules set in advance (not the focus of this paper, and will not be repeated). After completing federated learning, the final contribution index  $c^{(j)}$  of each participant is the output.

*Step 5.* The FL advertising platform charges the data providers for advertising fees  $p_j$ , in accordance with the payment rules.

**3.1.2. Mechanism Design Environment Setting.** In this paper, we propose an ad-sponsored search platform for federated learning. Aiming to solve the matching problem of participants in the federated learning scenario, our goal is to find a set of optimal allocation rules and payment rules to maximize the benefits of the FL advertising platform and meet certain constraints on the design of sponsored search auction mechanisms. Our assumptions were as follows:

- (1) There are  $n$  data owners  $j$  who are interested in a certain federated learning task, where  $J = \{1, 2, \dots, n\}$  represents a set of data providers. In addition, the alliance has  $m$  advertising slots  $k$ , where data providers can place their own basic data information (e.g., sample size, characteristics, data quality, and data cost), where  $K = \{1, 2, \dots, m\}$  denotes the set of these advertising slots.
- (2)  $\alpha_{jk}$  is the probability that a data user clicks when the data provider  $j$  is in the  $k^{\text{th}}$  advertisement slot, where



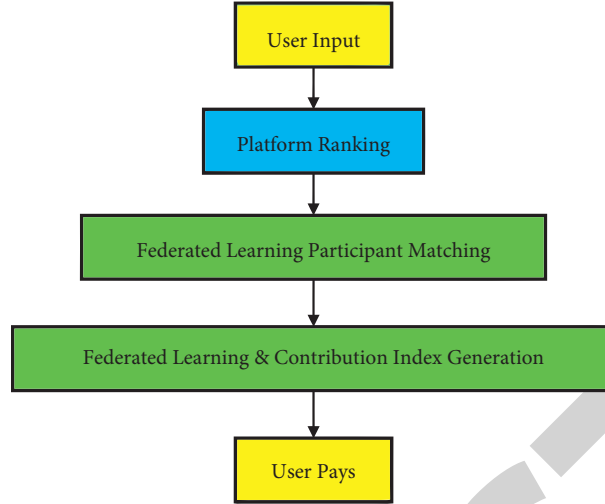


FIGURE 2: Schematic diagram of the advertising space of the FL advertising platform.

slot 1 represents the highest slot (the most prominent slot). We assume that  $\alpha_{jk}$  satisfies the following conditions:

$$1 \geq \alpha_{j1} \geq \alpha_{j2} \geq \dots \geq \alpha_{jm} \geq 0 \forall j \in J \forall k \in K. \quad (1)$$

- (3) For any federated learning task, each data provider  $j$  who is interested in this task makes a bid  $b^{(j)} \geq 0$ , where  $b^{(j)}$  depends on their economic behavior in the real market environment, with a certain degree of randomness.
- (4) Each data provider knows exactly how much value they can get when any data user clicks on their advertisement once. We assume that this value has nothing to do with the ad position, and this value only depends on whether the data user clicks on its ad. The data provider  $j$  does not know the value that other data providers get from one click of the data user. Formally, we assume that the value that data provider  $j$  gets from each click of a data user is  $v^{(j)}$ . The parameter  $v^{(j)}$  is called the valuation of data provider  $j$ , and the set of possible valuations of this provider is denoted as  $V^{(j)}$ .
- (5) Each data provider  $j$  has a federated learning contribution index  $c^{(j)}$ , which measures the quality of their own data. It is worth noting that the contribution degree represents the true contribution ability of the data provider; it has nothing to do with the bid and can truly reflect the value of the data owned by the data provider. This contribution can be calculated through a variety of mechanisms. In this paper,

we use the method proposed by Wei et al. [20] to calculate the contribution index based on the Shapley value.

- (6) Each data provider  $j$  is rational and intelligent, which means they pursue the maximum expected value of the utility function  $u^{(j)}$ , which we define later.

According to the above model assumptions, the sponsored search auction problem of the FL advertising platform can be accurately described as follows. Whenever the FL advertising platform receives the keywords of this federated learning task, it uses the bid profile  $b = (b^{(1)}, b^{(2)}, \dots, b^{(j)}, \dots, b^{(n)})$  and contribution index profile  $c = (c^{(1)}, c^{(2)}, \dots, c^{(j)}, \dots, c^{(n)})$  to determine: (1) Which data providers can win the slots to present their advertisements, as well as the order in which their advertisements are presented and (2) when data users click on their advertisements, the amount of money each data provider should pay to the platform. The former is the allocation rule, and the latter is the payment rule.

We define the problem environment of the sponsored search mechanism design in a typical FL advertising platform as follows:

- (i) Result set  $X$ : A result of the federated advertising sponsored search auction is a vector  $x = (y_{jk}, p_j)$ ,  $\forall j \in J; \forall k \in K$ , where  $y_{jk}$  is the probability that data provider  $j$  is assigned to slot  $k$ , and  $p_j$  is the cost-per-click paid by the data provider to the FL advertising platform. Therefore, the feasible result set can be expressed as:

$$X = \left\{ (y_{jk}, p_j) \mid y_{jk} \in \{0, 1\}; \sum_{j=1}^n y_{jk} \leq 1; \sum_{k=1}^m y_{jk} \leq 1; p_j \geq 0, \forall j \in J; \forall k \in K \right\}. \quad (2)$$

- (ii) Utility function  $u^{(j)}(\bullet)$ : Given a result  $x$ , the utility function of data provider  $j$  can be expressed as:

$$u^{(j)}(x(r), b^{(j)}) = \left( \sum_{k=1}^m y_{jk} \alpha_{jk} \right) (b^{(j)} - p_j). \quad (3)$$

- (iii) Social surplus function  $W$ : In this case, the utility of all data providers constitutes the social surplus function of the entire alliance. This is expressed as:

$$W = \sum_{j=1}^n u^{(j)}(x(r), b^{(j)}), \quad \forall j \in J. \quad (4)$$

For convenience, we summarize the defined symbols and their meanings in Table 1.

**3.2. Federated Learning Advertising Platform Sponsored Search Auction Mechanism Design.** We define the sponsored search mechanism design problem of a typical federated learning advertising platform as follows:

Let  $b = (b^{(1)}, b^{(2)}, \dots, b^{(j)}, \dots, b^{(n)})$  and  $c = (c^{(1)}, c^{(2)}, \dots, c^{(j)}, \dots, c^{(n)})$  be the bid profile and contribution index profile of data provider  $j$ , respectively. Let  $b^{(-j)}$  and  $c^{(-j)}$  denote the bid and contribution index profile of all other participants except for data provider  $j$ , respectively. As the bid profile is affected by multiple factors in actual production and the real world, it has a certain degree of randomness, which may lead to the problem of malicious bidding. For this reason, we propose a new ranking index,  $r = (r^{(1)}, r^{(2)}, \dots, r^{(n)})$ , which

TABLE 1: List of mathematical symbols and their definitions.

Symbol	Meaning
$j$	Index of the data provider
$k$	Index of the advertising slot
$\alpha_{jk}$	Probability that the data user clicks the $k^{\text{th}}$ ad slot
$b^{(j)}$	Bid from data provider $j$
$v^{(j)}$	Valuation from data provider $j$
$c^{(j)}$	FL contribution index of data provider $j$
$r^{(j)}$	Ranking score of data provider $j$
$x = (y_{jk}, p_j)$	Result set
$u^{(j)}(\bullet)$	Utility function of data provider $j$
$W$	Social surplus function

is composed of the bid profile  $b^{(j)}$  and the contribution index profile  $c^{(j)}$ , that is,

$$r^{(j)}(b^{(j)}) \stackrel{\text{def}}{=} (\beta \varphi^{(j)}(b^{(j)}) + (1 - \beta) c^{(j)}(\varphi^{(j)}(b^{(j)}))), \quad (5)$$

where  $c^{(j)}(\bullet)$  is contribution index function. In Ref. [20], Wei et al. proposed an effective federated learning contribution index calculation method based on the Shapley value. We also use a similar method to calculate the contribution index:

$$\begin{aligned} c^{(j)}(\varphi^{(j)}(b^{(j)})) \\ = \sum_{S \subseteq J \setminus \{j\}} \frac{|S|! (|J| - |S| - 1)!}{|J|!} \{v'(S \cup \{j\}) - v'(S)\}, \end{aligned} \quad (6)$$

where  $S$  is a subset of  $J$ ,

$$\begin{aligned} v'(S \cup \{j\}) &= \sum_{j' \in (S \cup \{j\})} \varphi^{(j')}(b^{(j')}) = \sum_{j' \in (S \cup \{j\})} \left( b^{(j')} - \frac{1 - F^{(j')}(b^{(j')})}{f^{(j')}(b^{(j')})} \right), \\ v'(S) &= \sum_{j' \in S} \varphi^{(j')}(b^{(j')}) = \sum_{j' \in S} \left( b^{(j')} - \frac{1 - F^{(j')}(b^{(j')})}{f^{(j')}(b^{(j')})} \right), \end{aligned} \quad (7)$$

where  $\varphi^{(j')}$  represents the virtual value of data provider  $j'$ ,

$$\varphi^{(j')}(b^{(j')}) = b^{(j')} - \frac{1 - F^{(j')}(b^{(j')})}{f^{(j')}(b^{(j')})}, \quad (8)$$

where  $F^{(j')}(b^{(j')})$  is the cumulative distribution of its bid, and the corresponding probability density function is  $f^{(j')}(b^{(j')})$ . These need to satisfy the following assumptions:

**Assumption 1** (Distribution of bid). For each data provider  $j$ , we assume that their bid  $b^{(j)}$  is independently—but not necessarily identically—drawn from a known cumulative distribution  $F^{(j)}(b^{(j)})$ , where the corresponding probability density function is  $f^{(j)}(b^{(j)})$ .

**Assumption 2** (Monotonicity of virtual value). For each data provider  $j$ , we assume that the cumulative distribution of its bid  $F^{(j)}(b^{(j)})$  satisfies the regular condition; thus, the virtual value  $\varphi^{(j)}(b^{(j)})$  is monotone nondecreasing.

Note that  $\beta \in [0, 1]$  is a variable parameter indicating the proportion of virtual value  $\varphi^{(j)}(\bullet)$  to the contribution index  $c^{(j)}(\bullet)$ . We provide its calculation method later.

After the FL advertising platform receives the participant's bids, the platform calculates the ranking score  $r$ , and then ranks the participants according to their ranking score. Therefore, a mechanism  $X = (x(r), p(r))$  consists of two rules, the allocation rule  $x(r)$  and the payment rule  $p(r)$ ; more specifically,  $x(r) = (x_1(r), x_2(r), \dots, x_n(r))$  and  $p(r) = (p_1(r), p_2(r), \dots, p_n(r))$ , where  $x_j(r) = \sum_{k=1}^m y_{jk} \alpha_{jk}$ ,  $\forall j \in J$ ;  $\forall k \in K$ ,  $y_{jk}$  is the probability that the data provider  $j$  is assigned to the  $k^{\text{th}}$  slot, and  $p_j$  is the cost-per-

click that the data provider pays to the FL advertising platform. Therefore, the feasible result set,  $X$ , can be expressed as:

$$X = \left\{ (y_{jk}, p_j) \mid \forall y_{jk} \in \{0, 1\}; \sum_{j=1}^n y_{jk} \leq 1; \sum_{k=1}^m y_{jk} \leq 1; p_j \geq 0, \forall j \in J; \forall k \in K \right\}. \quad (9)$$

It is worth noting that random results are also included in the above result set, which means that random mechanisms are also part of the mechanism space.

Given the result  $x$ , the utility function  $u^{(j)}(\bullet)$  of the data provider  $j$  can be expressed as:

$$u^{(j)}(x(r), b^{(j)}) = \left( \sum_{k=1}^m y_{jk} \alpha_{jk} \right) (b^{(j)} - p_j). \quad (10)$$

In the FL advertising platform system, to ensure that data providers always have the willingness to auction, we require that the utility of data provider  $j$  should not be less than zero, that is, we demand that each data provider satisfies individual rationality. We define this as follows:

*Definition 1* (Individual Rationality).

$$u^{(j)}(x(r), b^{(j)}) \geq 0, \quad \forall j \in J; \forall k \in K. \quad (11)$$

In addition, to avoid malicious bidding by data providers, we impose Bayesian Incentive Compatibility (BIC) constraints on the designed mechanism:

*Definition 2* (Bayesian Incentive Compatibility).

$$u^{(j)}(x(r), v^{(j)}) \geq u^{(j)}(x(r), b^{(j)}), \quad \forall j \in J. \quad (12)$$

If and only if a mechanism meets the conditions of IR and BIC, we call it a feasible mechanism. Myerson's theorem gives equivalent expressions for IR and BIC. Similarly, we can also design a mechanism with IR and BIC on the FL advertising platform, as shown below:

**Lemma 1** (Myerson theorem) [37]. *If and only if a mechanism satisfies IR and BIC, for any advertisement item from*

*data provider  $j$  and bids of other items  $b^{(-j)}$ , the allocation rule  $y_j(b^{(j)}, b^{(-j)})$ , is monotone nondecreasing on  $b^{(j)}$ , and the payment rule is as follows:*

$$p_j(b) = \begin{cases} b^{(j)} - \frac{\int_0^{b^{(j)}} y_j(x^{(j)}, b^{(-j)}) dx^{(j)}}{y_j(b)}, & y_j(b) \neq 0. \\ 0, & \text{else.} \end{cases} \quad (13)$$

In summary, we have the following mechanisms.

(1) Allocation rules

$$y_{jk}(b) = \begin{cases} 1, & \text{if } r^{(j)}(b^{(j)}) \text{ is } k^{\text{th}} \text{ highest ranking score, } k \in K, \\ 0, & \text{else.} \end{cases} \quad (14)$$

(2) Payment rules

$$p_j(b) = \begin{cases} b^{(j)} - \frac{\int_0^{b^{(j)}} y_j(x^{(j)}, b^{(-j)}) dx^{(j)}}{y_j(b)}, & y_j(b) \neq 0. \\ 0, & \text{else.} \end{cases} \quad (15)$$

Now, we give the calculation method for the coefficient  $\beta$ . Specifically, in such a mechanism, we require that the total social utility should be maximized, that is, the total social surplus  $W$  is maximized under certain constraints, making the malicious bidding rate  $\mu$  (defined later) lower than that in classic sponsored search auction mechanisms.

$$\max W. \quad (16)$$

$$\text{s.t.} \begin{cases} u^{(j)}(x(r), b^{(j)}) \geq 0; \\ u^{(j)}(x(r), v^{(j)}) \geq u^{(j)}(x(r), b^{(j)}); \\ \sum_{k=1}^m y_{jk} \leq 1; \\ \sum_{j=1}^n y_{jk} \leq 1; \\ 0 \leq y_{jk} \leq 1; \\ p_j \geq 0; \\ \forall j \in J; \\ \forall k \in K. \end{cases} \quad (17)$$

For the above constrained programming, we use the Lagrange multiplier method to solve the equation. The Lagrange function can be expressed as follows:

$$L = W + \left( \lambda \sum_{j=1}^n u^{(j)}(x(r), b^{(j)}) + \sum_{j=1}^n (u^{(j)}(x(r), v^{(j)}) - u^{(j)}(x(r), b^{(j)})) \right. \\ \left. + \sum_{j=1}^n \left( 1 - \sum_{k=1}^m y_{jk} \right) + \sum_{k=1}^m \left( 1 - \sum_{j=1}^n y_{jk} \right) + \sum_{j=1}^n \sum_{k=1}^m (1 - y_{jk}) + \sum_{j=1}^n \sum_{k=1}^m y_{jk} + \sum_{j=1}^n p_j \right), \quad (18)$$

where  $\lambda$  is the Lagrange multiplier,  $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_{4n+m+2mn})$ .

For ease of the descriptions, we use  $\Omega$  to represent all the constraints. Thus, the Lagrange function can be expressed as follows:

$$L = W + \lambda \cdot \Omega, \quad (19)$$

where  $\Omega \geq 0$ .

As  $W$  is a linear function of  $y$ , the constraints are also linear in  $y$ , and the result set is a convex set, and the constrained program in (16) and (17) consists of convex optimization. We can obtain the optimal value of the coefficient  $\beta$  in the equation of ranking score by Lemma 2.

**Lemma 2** (Optimal estimation of coefficient) [45]. *ie constrained program with constraints given in (17) is equivalent to the unconstrained programming with coefficient  $\beta^*$  and, when the optimal Lagrange multiplier is  $\lambda^*$ , the coefficient can be expressed as  $\beta^* = 1/(\|\lambda^*\| + 1)$ . We obtain  $\|\lambda^*\|$  by Algorithm 1. Then, we can obtain  $\beta^*$ .*

Algorithm 1 shows the detailed process for calculating the ranking score  $r^{(j)}$  of each data provider  $j$ . Specifically, lines 1–12 are the calculation of coefficient  $\beta$ , while lines 13 and 14 are the calculation of the ranking score  $r^{(j)}$ .

**Theorem 1** (FSSA). *When a federated learning advertising platform satisfies Assumptions 1 and 2 under the conditions of Definitions 1 and 2, and takes Equation (5) as the ranking score, an optimal mechanism can be obtained.*

*Proof.* When Assumptions 1 and 2 are satisfied, we can construct a ranking score based on contribution index and virtual value with the help of Equations (6) and equation (8), that is, Equation (5). As it is under the conditions of Definitions 1 and 2, Lemma 1 holds, and the calculation expressions of allocation rules and payment rules can be obtained. As the linear program constructed under the above conditions satisfies convexity, Lemma 2 holds, and the optimal estimate of the coefficients in the ranking score can be obtained. Q.E.D.

**3.3. Simulation Experiment Design.** To evaluate the performance of our proposed federated learning advertising platform sponsored search auction mechanism (FSSA)

algorithm, we carried out many simulation experiments using *Python*. To verify whether the malicious bidding rate of FSSA had changed, we compared it with four classic mechanisms based on bid  $b_i$  ranking, under the same simulation data set and experimental environment settings. The experiment was run on a Windows 10 desktop with 32 GB main memory, an Intel Xeon E5-2690 v3 @2.60 GHz(X2) CPU with 12 cores and 24 threads, and an NVIDIA GeForce GTX1060 6 GB graphics card.

**3.3.1. Data Set.** In a real market environment, a data provider's bid is related to many factors, and the data provider may not report its bid  $b^{(j)}$  in a truthful way. Therefore, in this experiment, we assumed that each bid  $b^{(j)}$  obeys a uniform distribution, that is,  $b^{(j)} \sim U(0, 1)$ . Moreover, as the number of data providers and the number of advertising slots cannot be determined, we divided it into two situations to conduct separate experiments: (1) When  $n > m$ , assume that the number of data providers  $n$  is 8, and the number of ad slots  $m$  is 5, 6, or 7. The method of setting the click-through rate obeys the following principle: the click-through rate increases at equal intervals from 0.1 to 0.9, according to the number of ad slots (e.g., when  $m = 5$ , the click probability  $\alpha_{jk}$  of each advertising spot, from bottom to top, is from 0.1 to 0.9, having the same interval of 0.2 increases successively) and (2) when  $n \leq m$ , assume that the number of data providers  $n$  is 5, and the number of ad slots is  $m$  and the click probability  $\alpha_{jk}$  are the same as when  $n > m$ .

Algorithm 2 shows the detailed process for generating the training data set. Specifically, line 1 is the process of generating all possible ranking results set. As for each sample in the training data set, lines 3–17 show its generation process. Lines 18–20 show the process of generating a label for each training data sample, initializing the set  $\{r^{(j)}\}$  to set  $\{r_{\text{order}}^{(j)}\}$  first, then ranking the elements in it in the descending order, using  $r_{\text{order}} = (r_{\text{order}}^{(1)}, r_{\text{order}}^{(2)}, \dots, r_{\text{order}}^{(j)}, \dots, r_{\text{order}}^{(n)}) \in \{\text{list}_l | l = 1, 2, \dots, (\min(m, n))!\}$  to represent the ranking result. Finally,  $r_{\text{order}}$  is matched in the set  $\{\text{list}_l\}$ , and the index of the matched  $\text{list}_l$  is taken as the label of the training data sample.

**3.3.2. Comparison Algorithm.** We compared the malicious bidding rate of the proposed method with those of three classic sponsored search auction algorithms, which have been summarized by Narahari et al. in Ref. [3]. Due to space

Input : bid  $b^{(j)}$  of each data provider  $j$ ; cumulative distribution  $F^{(j)}(b^{(j)})$ ; probability density function  $f^{(j)}(b^{(j)})$ ; any positive number  $\delta$ ; initial upper bound  $\|\lambda\|_{up}$  which makes  $\Omega \geq 0$ ; lower bound  $\|\lambda\|_{low}$ .

Output :  $b^{(j)}$ .

- (1) Initialize  $\|\lambda\|$  to 0 and  $\beta$  to 1 respectively.
- (2) if  $\Omega \geq 0$  then
- (3)  $\|\lambda^*\| \leftarrow 0$ .
- (4) else
- (5)  $\|\lambda\|_{low} \leftarrow 0$ .
- (6) while  $\|\lambda\|_{up} - \|\lambda\|_{low} > \delta$  do:
- (7)  $\|\lambda\|_{temp} \leftarrow \text{mean}(\|\lambda\|_{up}, \|\lambda\|_{low})$ .
- (8) Assign  $\|\lambda\|_{temp}$  to  $\|\lambda\|_{up}$  when it makes  $\Omega \geq 0$ , otherwise assign it to  $\|\lambda\|_{low}$ .
- (9) end while:
- (10)  $\|\lambda^*\| \leftarrow \|\lambda\|_{temp}$ .
- (11) end if
- (12)  $\beta^* \leftarrow 1/(\|\lambda^*\| + 1)$ .
- (13) Use Equations (1 – 3) to obtain  $r^{(j)}$ .
- (14) return  $r^{(j)}$

ALGORITHM 1: Calculation of the ranking score  $r^{(j)}$ .

Input : ranking score  $r^{(j)}$  of each data provider  $j$ , bid  $b^{(j)}$  of each data provider  $j$ ,  $\forall j \in J$ , training data set size  $N$ .

Output : training data set  $\{x_i\}$  where  $x_i = (y_{jk}, p_j)$ ,  $\forall j \in J, \forall k \in K$ ; label label <sub>$i$</sub>  of each training data sample,  $i = 1, 2, \dots, N$ .

- (1) Generate all possible ranking results set  $\{\text{list}_l | l = 1, 2, \dots, (\min(m, n)!) \}$  of set  $\{r^{(j)} | \forall j \in J\}$  by full permutation operation.
- (2) for each  $i \in \text{range}(N)$  do
- (3) for each  $j \in \text{range}(n)$  do
- (4) for each  $k \in \text{range}(m)$  do
- (5) if  $r^{(j)}(b^{(j)})$  is the  $k^{\text{th}}$  highest ranking score then
- (6)  $y_{jk} \leftarrow 1$ .
- (7) else
- (8)  $y_{jk} \leftarrow 0$ .
- (9) end if
- (10) end for
- (11) if  $y_j = 1$  then
- (12)  $p_j \leftarrow b^{(j)} - \int_0^{b^{(j)}} y_j(x^{(j)}, b^{(-j)}) dx^{(j)} / y_j(b)$ .
- (13) else
- (14)  $p_j \leftarrow 0$ .
- (15) endif
- (16) end for
- (17)  $x_i \leftarrow (y_{jk}, p_j)_i$ .
- (18) Initialize  $\{r^{(j)}\}$  to  $\{r_{\text{order}}^{(j)}\}$ .
- (19) Rank the elements in  $\{r_{\text{order}}^{(j)}\}$  in descending order.
- (20) Match  $\{r_{\text{order}}^{(j)}\}$  with each element in set  $\{\text{list}_l\}$  to obtain label <sub>$i$</sub> .
- (21) return  $x_i, \text{label}_i$ .
- (22) end for

ALGORITHM 2: Training data set generation.

limitations, we do not review these algorithms in detail; instead, brief summaries are provided in the following.

**GFP.** This algorithm is the generalized first price auction mechanism. Under this mechanism, the  $m$  advertising positions are allocated to advertisers in the descending order, according to their bids.

**GSP.** This algorithm is the generalized second-price auction mechanism, which is adapted from the generalized first price auction mechanism. Due to the instability of the generalized first price, the entire system has low

allocation efficiency; thus, GSP was proposed. This mechanism is completely consistent with the GFP allocation rules, but changes its payment rules. The participant who gets slot 1 pays the bid of the participant who wins slot 2, and so on.

**VCG.** This algorithm is the abbreviation of the Vickery–Clark–Groves mechanism, which is the most widely used mechanism in the field of mechanism design. In a quasilinear environment, the VCG mechanism not only has allocation efficiency but also incentive compatibility.



**3.3.3. Evaluation Metrics.** To evaluate whether our proposed algorithm can effectively reduce the probability of malicious bidding problems or not, we defined a new evaluation index, the malicious bidding rate  $\mu$ , which represents the cumulative number of malicious bidding events happening in  $N$  bidding scenarios.

$$\mu \stackrel{\text{def}}{=} \frac{1}{N} \sum_{i=1}^N M, \quad (20)$$

where  $M$  represents a malicious bidding event. We define the distance between the result of the bid  $b$  or the ranking score  $r$  and the benchmark ranking  $c$  as the similarity  $d$ . If the similarity  $d$  is not 0, it indicates that a malicious bidding event  $M$  has happened, that is,

$$d(r_{(.)}, r_{(c)}) \stackrel{\text{def}}{=} \begin{cases} \frac{1}{\text{score}(r_{(.)}, r_{(c)})}, & \text{if } \text{score}(r_{(.)}, r_{(c)}) \neq 0, \\ +\infty, & \text{else,} \end{cases} \quad (21)$$

where  $r_{(.)}$  represents the result of ranking by bid  $b$  or ranking score  $r$ ;  $r_{(c)}$  represents the result of ranking by the contribution index  $c$ , which we call benchmark ranking; and  $\text{score}(r_{(.)}, r_{(c)})$  represents the matching score between the ranking results. That is, for two sequences with the same length of  $m$ , starting from the first slot, if the  $k^{\text{th}}$  slot of  $r_{(.)}$  and  $r_{(c)}$  is the same, then the match of this slot is successful, and 1 point is counted; otherwise, 0 points are counted. Formally, we define the rank score as  $\text{score}(r_{(.)}, r_{(c)}) \stackrel{\text{def}}{=} \sum_{k=1}^m \text{Match}(r_{(.)}, r_{(c)})$ , where  $\text{Match}(r_{(.)}, r_{(c)})$  represents the matching situation, which is:

$$\text{Match}(r_{(.)}, r_{(c)}) \stackrel{\text{def}}{=} \begin{cases} 1, & \text{if } r_{(.)}^{(k)} = r_{(c)}^{(k)}, \quad \forall k \in K. \\ 0, & \text{else,} \end{cases} \quad (22)$$

Algorithm 3 shows the calculation process for the malicious bidding rate. In lines 1–7, we use Equation (22) to calculate the matching score and, in line 8, we use Equation (21) to calculate the similarity. Finally, we use Equation (20) to calculate the evaluation index malicious bidding rate  $\mu$  in lines 9 and 10.

**3.3.4. Experimental Setup.** In this paper, a convolutional neural network (CNN) model was used for training. Algorithm 4 shows the training process of the convolutional neural network (CNN) model, using the training data set and label set generated in Algorithm 2. Lines 1–3 show the initialization of the model parameters and the setting of the learning rate, while Lines 4–10 show the process of obtaining the optimal parameters of the neural network model through the backward propagation algorithm. Specifically, the model consists of two convolution layers and three fully connected

layers. The output of each hidden layer is activated by the ReLU activation function, the dropout is set to 0.5. The backward propagation algorithm is used to solve the parameters, the optimizer is set to Adam, the loss function is set to sparse categorical crossentropy, and the learning rate is set to 0.001. The training batch size is set to 124, and the number of iterations is 200 rounds.

**3.3.5. Analysis of Results.** As shown in Figure 3, the experimental results under different simulation experimental data sets confirmed that the performance of the FSSA mechanism proposed in this paper was better than that of the other benchmark mechanisms. This further proves the effectiveness and robustness of the FSSA mechanism for the ranking of the advertising slots in federated learning advertising platforms.

Specifically, when  $n \leq m$  (i.e.,  $n = 5$ ), from the horizontal perspective,  $m$  continues to increase. The accuracy of the four mechanisms showed little change. In this case, the number of participants was always less than the number of advertisements, so the participants were always fully allocated, and the complexity of the model changes mainly depend on  $\min(n, m)$ . Therefore, even if the number of advertising slots continues to increase, the accuracy of the model does not change much, and the model is relatively stable.

When  $n > m$  (i.e.,  $n = 8$ ), from a horizontal perspective, with the continuous increase of  $m$ , the accuracy of the four mechanisms gradually decreased, and the speed of convergence also decreased. In this case, the number of participants was always greater than the number of advertisements, so there was the situation where there were remaining participants. The complexity of the model mainly depends on  $\min(n, m)$ . Therefore, with the continuous increase of  $m$ , the allocation rules that need to be considered are more complicated and the model complexity is high. In the end, the model converges slowly and the accuracy of the convergence is also reduced, in accordance with the actual situation.

For each case of  $m$ , with  $m$  is fixed, vertical comparison between  $n = 5$  and  $n = 8$  indicated that, when  $n = 5$ , the number of participants is less than the number of advertising slots, and the advertising slots cannot be fully allocated. Compared with  $n = 8$ , the latter had a richer allocation and a higher allocation efficiency, thus facilitating model convergence.

In Table 2, our testing accuracy of the model in different situations is shown. It can be seen that, no matter the situation, the testing accuracy of our proposed FSSA mechanism was higher than that of the benchmark.

As can be seen from Figure 4, when  $n = 5$ , that is, participants were always less than the number of advertisements. With the increase of  $m$ , the accuracy of the four mechanisms showed little change, and this is because the remaining advertising space is always sufficient. For each participant, they are not sensitive to the dynamic changes of



```

Input : label set  $\{\text{label}_{r_i}\}$  based on  $r$ ; label set  $\{\text{label}_{c_i}\}$  based on  $c$ ,  $i = 1, 2, \dots, N$ ; label set  $\text{label}_i$ ,  $i = 1, 2, \dots, N$ ; all possible ranking results set  $\text{label}_i$ ,  $i = 1, 2, \dots, (\min(m, n))!$ 
Output : malicious bidding rate  $\mu$ .
(1) for each  $i \in \text{range}(N)$  do
(2)   for each  $j \in \text{range}(n)$  do
(3)     if  $\text{list}_{\text{label}_{r_i}}^{(j)} = \text{list}_{\text{label}_{c_i}}^{(j)}$  then
(4)       score  $\leftarrow$  score + 1.
(5)     end if
(6)   end for:
(7) end for
(8) Take the reciprocal as measurement and record the independent cases by count.
(9)  $\mu \leftarrow (\text{count}/N)$ .
(10) return  $\mu$ .

```

ALGORITHM 3: Calculation of the malicious bidding rate  $\mu$ .

```

Input : training data set  $\{x_i\}$ , label set  $\text{label}_i$ ,  $i = 1, 2, \dots, N$ .
Output : model parameters  $\omega, \tau$ , where  $\omega$  is weight coefficient and  $\tau$  is bias coefficient.
(1)  $PAR \leftarrow$  list of all parameters.
(2) Initialize  $PAR$ .
(3) Set learning rate to  $\eta$ .
(4) for each  $t \in \text{range}(\text{iterations})$  do
(5)   Set the loss function to Loss.
(6)   Feed Forward (Loss).
(7)   Backward Propagation (Loss).
(8)    $\delta = d\text{Loss}/dPAR$ 
(9)    $PAR \leftarrow PAR - \eta * \delta$ .
(10) end for
(11) return  $\omega, \tau$ .

```

ALGORITHM 4: CNN Training.

allocation rules or payment rules in the process of competition. Therefore, the learning degree of the model to the data features contained in the sample data tensor tends to be saturated. And, in this scenario, the testing accuracy of our proposed FSSA mechanism was higher than others. This means that FSSA can still learn the data distribution law relatively deeply when the data features are not obvious, so as to estimate the model parameters more accurately.

As can be seen from Figure 5, when  $n = 8$ , that is, participants were always more than the number of advertisements. With the continuous increase of  $m$ , the testing accuracy of various mechanisms varies. And, in this scenario, there were remaining participants, and the complexity of the model mainly depends on  $\min(n, m)$ . Due to the fierce competition, the data tensor contains rich features, but it is also easy to cause the phenomenon of over fitting of model parameters, which makes the generalization ability of the model insufficient. In this case, GFP mechanism and GSP mechanism are difficult to deal with such problems, while VCG mechanism and FSSA mechanism alleviate the over fitting phenomenon on the premise of ensuring relatively high accuracy, and FSSA mechanism can still maintain relatively high testing accuracy. Thus, the feasibility and stability of the model are trustworthy.

We summarize the malicious bidding rates in different situations in Table 3. It can be seen, from the table, that the malicious bidding rate under our proposed FSSA mechanism was as low as 0.01016. Compared with the classic sponsored search auction mechanisms, the malicious bidding rate was greatly decreased under the same conditions, effectively reducing the probability of malicious bidding in the auction process for advertising slots on the Federated Learning advertising platform.

In addition, it can be seen, from Figure 6, that when  $n = 5$ ,  $m$  increased, as all participants were allocated, and they had no motivation to lie. Therefore, the trend of the malicious bidding rate for the four mechanisms was not obvious. Within the same group of  $m$  and  $n$ , the malicious bidding rate under the conditions of GFP mechanism and GSP mechanism is relatively high. Therefore, it can be seen that the response ability of these two mechanisms to the phenomenon of malicious bidding is insufficient. In contrast, VCG mechanism reduces the malicious bidding rate to about 0.077, which is relatively obvious. Even so, the malicious bidding rate of our proposed FSSA mechanism was still far lower than other benchmark mechanisms, and it is always lower than 0.04. As shown in the figure, this mechanism has a significant effect on reducing the malicious bidding rate.

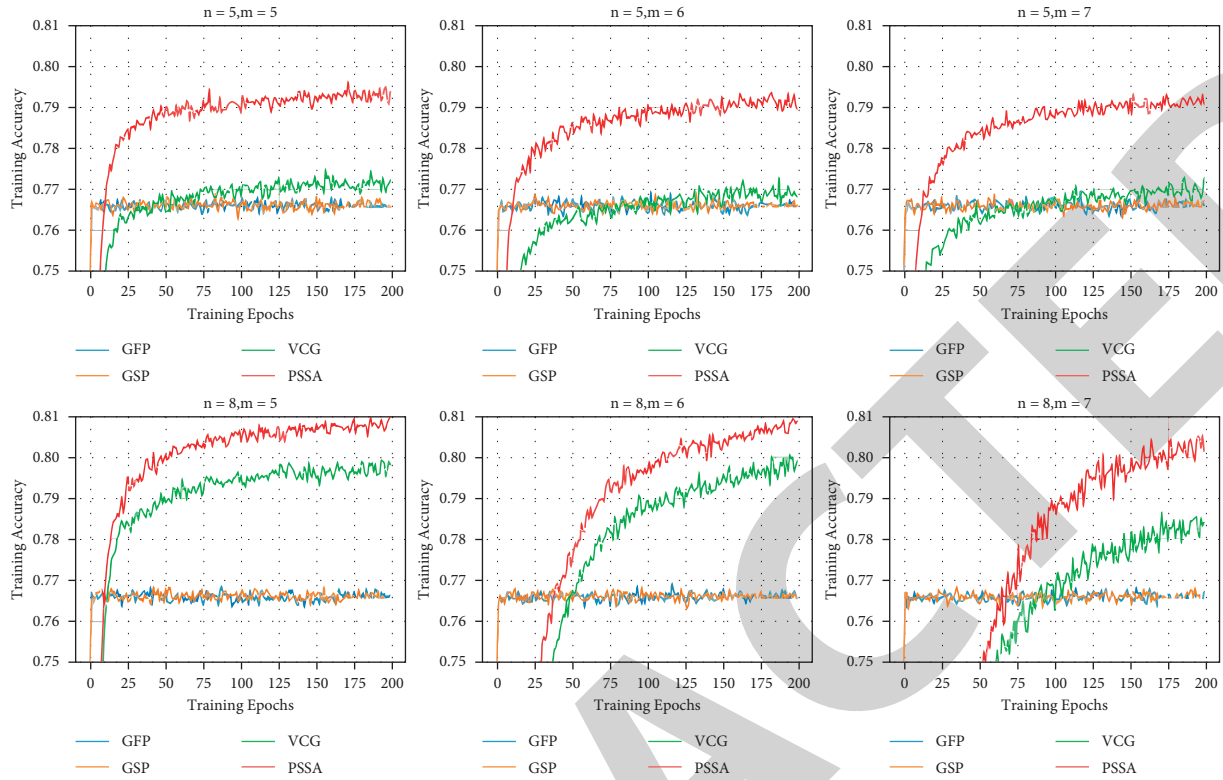


FIGURE 3: Training accuracy of different mechanisms in various situations.

TABLE 2: Testing accuracy of different mechanisms in various situations.

(n,m)	GFP	GSP	VCG	FSSA
(5,5)	0.76434	0.76551	0.77223	0.78366
(5,6)	0.76509	0.76601	0.77206	0.78949
(5,7)	0.76584	0.76706	0.77270	0.79063
(8,5)	0.76634	0.76449	0.79811	0.80984
(8,6)	0.76676	0.76427	0.79907	0.80877
(8,7)	0.76716	0.76494	0.78411	0.80143

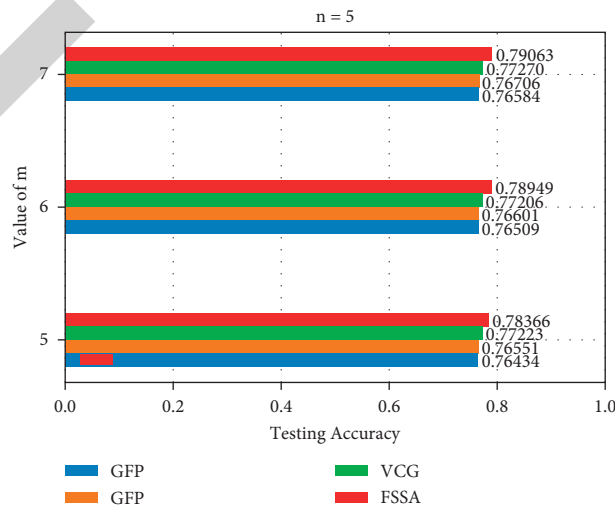


FIGURE 4: Testing accuracy of different mechanisms at  $n = 5$ .

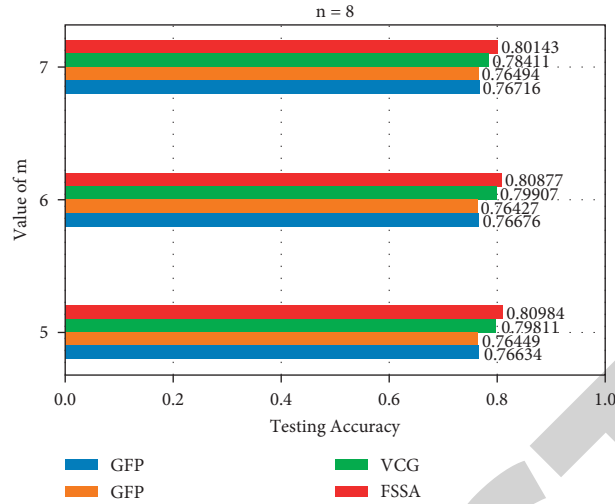


FIGURE 5: Testing accuracy of different mechanisms at  $n = 8$ .

TABLE 3: Malicious bidding rate of different mechanisms in various situations.

(n,m)	GFP	GSP	VCG	FSSA
(5,5)	0.12566	0.11449	0.07777	0.03634
(5,6)	0.12491	0.11399	0.07794	0.03051
(5,7)	0.12416	0.11294	0.07730	0.02937
(8,5)	0.22366	0.21551	0.15199	0.01016
(8,6)	0.22324	0.21573	0.15093	0.01123
(8,7)	0.22284	0.21506	0.16589	0.01857

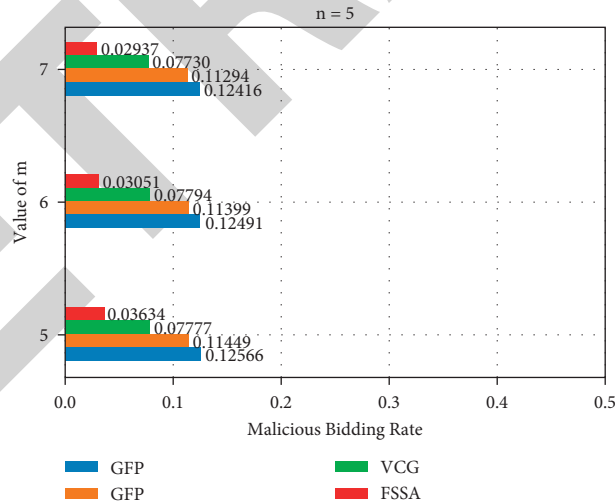


FIGURE 6: Malicious bidding rate under different mechanisms at  $n = 5$ .

However, Figure 7 shows that, at  $n = 8$ , due to the surplus of participants, participants had greater motivation to lie. Therefore, when  $m$  increase linearly, the rate of malicious bidding also gradually increased; however, that of the proposed mechanism was still lower than those of the other benchmark mechanisms, in line with our expectations. Moreover, when we compare the experimental results shown in Figure 7 with those shown in Figure 6, we find that when  $m$  takes the same value, that is, when the same number of

advertisements are supplied, the competition among participants is fierce, that is, as shown in Figure 7, the response ability of GFP, GSP, and VCG to malicious bidding is weakened, and only FSSA has not received any negative impact. Compared with the experimental results shown in Figure 6, its response ability to malicious bidding rate is more significant, and the malicious bidding rate is further reduced.

In summary, compared with other benchmark sponsored search auction mechanisms, our proposed FSSA

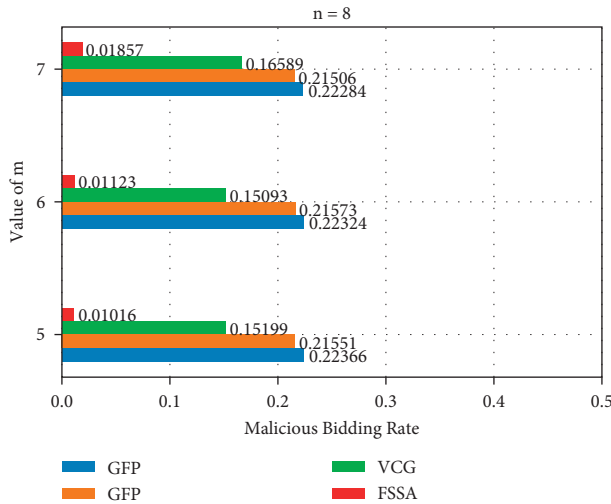


FIGURE 7: Malicious bidding rate under different mechanisms at  $n = 8$ .

mechanism can effectively curb the generation of malicious bidding in the process of sponsored search auctions for federated learning advertising slots without imposing incentive-compatible constraints, which still allows for the social resources to be fully allocated.

#### 4. Conclusions

In this paper, we proposed a novel Federated Sponsored Search Auction Mechanism (FSSA) for the scenario of federated learning. Based on the original bidding, the average contribution index of the data providers in their previous federated learning scenarios is considered, and a certain weight is assigned to them. A new ranking index is formally defined, that we call ranking scores, which make the proposed mechanism more in line with the scenario of federated learning. We conducted a large number of experiments on different simulation data sets to verify the effectiveness of our proposed method. The experimental results demonstrated that our method can rank data providers fairly. Compared with classic sponsored search auction mechanisms, the malicious bidding rate was reduced significantly through the use of the proposed mechanism. The social resources were effectively allocated, and the establishment of the federated learning ecology was promoted.

However, when studying slot allocation in an FL advertising platform, we only considered the situation where there are a small amount of data providers and advertising slots. When the number of participants continues to increase, the stability of the model will decrease and the generalization ability of the model needs to be improved. In addition, due to the experimental cost and the requirements of data privacy protection regulations, we did not obtain a real experimental data set and only conducted our experiments on a simulated experimental data set. Our future work will focus on optimizing the parameters in order to improve the generalization ability of the used neural network models as well as to obtain real experimental data.

#### Data Availability

The data used to support the results of this study are currently in the privacy protection stage and can be obtained from the corresponding author upon request 6 months after the article is published.

#### Conflicts of Interest

The authors declare that they have no conflicts of interest.

#### Acknowledgments

This paper was funded by Shenzhen Educational Science Planning Project (7020320030), Guangdong Province Philosophy and Social Science Planning Project (GD15XGL52).

#### References

- [1] T. S. Qaid, H. Mazaar, M. Y. H. Al-Shamri, M. S. Alqahtani, A. A. Raweh, and W. Alakwaa, "Hybrid deep-learning and machine-learning models for predicting COVID-19," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 9996737, 11 pages, 2021.
- [2] G. Papaioannou and I. Sarakinos, *The General Data protection Regulation (GDPR, 2016/679/EE) and the (Big) Personal Data in Cultural Institutions: Thoughts on the GDPR Compliance Process*, Springer, Berlin, Germany, 2018.
- [3] Y. Narahari, D. Garg, R. Narayanam, and H. Prakash, "Game theoretic problems in network economics and mechanism design solutions," in *Advanced Information and Knowledge Processing*, Springer, Berlin, Germany, 2009.
- [4] P. Kairouz, H. B. McMahan, B. Avent et al., "Advances and open problems in federated learning," *Found. Trends Mach. Learn.*, vol. 14, pp. 1–210, 2021.
- [5] K. Hu, Y. Li, M. Xia, J. Wu, and M. Lu, "Federated learning: a distributed shared machine learning method," *Complexity*, vol. 2021, Article ID 8261663, 20 pages, 2021.
- [6] J. Konečný, H. B. McMahan, D. Ramage, and P. Richtárik, "Federated optimization: distributed machine learning for on-device intelligence," 2016, <https://arxiv.org/abs/1610.02527>.
- [7] Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated machine learning: concept and applications," *ACM Trans. Intell. Syst. Technol.*, vol. 10, pp. 1–19, 2019.
- [8] K. A. Bonaawitz, V. Ivanov, B. Kreuter et al., "Practical secure aggregation for privacy preserving machine learning," *International Association for Cryptologic Research*, vol. 281, 2017.
- [9] P. Mohassel and Y. Zhang, "SecureML: a system for scalable privacy-preserving machine learning," in *Proceedings of the IEEE Symp. Secur. Privacy*, pp. 19–38, San Jose, CA, USA, May 2017.
- [10] D. Gao, Y. Liu, A. Huang, C. Ju, H. Yu, and Q. Yang, "Privacy-preserving heterogeneous federated transfer learning," in *Proceedings of the 2019 IEEE International Conference on Big Data (Big Data)*, pp. 2552–2559, Los Angeles, CA, USA, 2019.
- [11] L. Song, C. Ma, G. Zhang, and Y. Zhang, "Privacy-Preserving unsupervised domain adaptation in federated setting," *IEEE Access*, vol. 8, Article ID 143233, 2020.
- [12] I. Giacomelli, S. Jha, M. Joye, C. D. Page, and K. Yoon, "Privacy-Preserving ridge regression with only linearly-homomorphic encryption," in *ACNS*, Springer, Berlin, Germany, 2018.

## Retraction

# Retracted: Dynamic Mechanism and Decision Analysis of Urban Exhibition Tourism Development Based on High-Performance Computing

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] H. Wan, "Dynamic Mechanism and Decision Analysis of Urban Exhibition Tourism Development Based on High-Performance Computing," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3806752, 14 pages, 2022.

## Research Article

# Dynamic Mechanism and Decision Analysis of Urban Exhibition Tourism Development Based on High-Performance Computing

Hongzhen Wan 

*School of Management, Guangdong Industry Polytechnic, Guangzhou 510300, China*

Correspondence should be addressed to Hongzhen Wan; 2009110035@gdip.edu.cn

Received 28 January 2022; Revised 1 March 2022; Accepted 15 March 2022; Published 6 April 2022

Academic Editor: Daqing Gong

Copyright © 2022 Hongzhen Wan. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

MICE tourism has received the attention and support from relevant departments and tourism industry with the growth rate of recent years. In this paper, the dynamics mechanism of urban exhibition tourism development is studied by using high-performance computing, and the dynamics mechanism model of exhibition tourism is constructed. To improve the performance, the solution constructed in the adaptive joint algorithm is used as the initial solution, and then the initial solution is improved using the method proposed by the new neighborhood strategy. It promotes the development of MICE tourism and urban MICE tourism by combining research on theoretical aspects related to tourism development, such as the relationship between tourism development and exhibition, and the development dynamics and stages of MICE tourism. The research results show that, in the assessment of the development mechanism of urban MICE tourism, tourism resources, economic, and policy factors have a relatively large weight, among which the weight coefficient of tourism resources factor is 0.2411. It can be seen that the development power of MICE tourism should start from tourism resources, so that the development power mechanism of urban MICE tourism can be optimized and improved, which not only provides new methods and new ideas for the transformation and upgrading of tourism industry but also promotes tourism enterprises and related departments to creatively develop tourism resources and design competitive tourism products, so as to improve the supply quality of MICE tourism, and then continuously promote the process of supply-side reform of MICE tourism.

## 1. Introduction

Scheduling and scheduling tools for high-performance computing (HPC) machines have a key role in mapping jobs to available resources, trying to maximize performance and quality of service (QoS). Bridi et al. demonstrated the first search-based HPC machine job allocation and scheduling method, working in a production environment. The scheduler is based on constraint programming, which is an effective programming technique for optimization problems. Tests on various operating conditions show that, in midlevel HPC machines equipped with the most advanced business rules-based scheduler, waiting and QoS have been significantly improved. In addition, Bridi et al. analyzed the conditions under which their method was superior to commercial methods to create a combination of scheduling algorithms that ensure robustness, flexibility, and scalability. Although their method has certain reference value for the

study of this article, it is of little value [1]. Computational models can easily promote the mechanical understanding of multiscale biological processes, such as cell proliferation in constantly changing biological tissues. Although there are tools for constructing and simulating multiscale models, the statistical inference of unknown model parameters is still an open question. Jagiella et al. demonstrated and benchmarked the Parallel Approximate Bayesian Computing Sequential Monte Carlo (pABC SMC) algorithm, which is tailored for high-performance computing clusters. Their research proved the principle of robust data-driven modeling of multiscale biological systems, and the feasibility of multiscale model parameterization through statistical inference. The algorithm is too complicated, and errors may occur in the calculation [2]. Wang et al. discussed the design trade-offs of extremely scalable system software. They proposed a general system software classification method by deconstructing general HPC system software into its basic



components. They provided a system method for building scalable system software by decomposing it into basic components, allowing system software to be classified according to the characteristics of these components for design evaluation through simulation or actual implementation. The results show that it helps to lay the foundation for the development of next-generation HPC system software for extreme scale. But this research is still in its infancy, and it is too basic [3]. The high-performance computing (HPC) community has spent decades developing tools and teaching practitioners to use the power of parallel and distributed computing. In order to create a scalable and flexible educational experience for practitioners at all stages of their careers, Mullen turned to Massive Open Online Courses (MOOC). He introduced in detail the design of a unique self-paced online course, and the course materials are presented through the lens of common HPC use cases and parallelization strategies. Using a personalized path, he teaches researchers how to identify the consistency between scientific applications and traditional HPC use cases so that they can focus on learning parallelization strategies that are critical to their workplace success. However, the high-performance computing method he proposed is not specific enough to be discussed in depth [4]. Fu's et al. research examines the potential dimensions of place attachment in the exhibition environment and its impact on participant satisfaction. The study took the China (Shenzhen) International Cultural Industry Expo as a case and proposed and tested a series of multiple intermediary models, divided into two groups: first-time participants and repeat participants. The exhibition attachment determines a function-emotion-identity sequence, examining direct and mediating effects and multiple group comparisons. Although most hypotheses have been confirmed, emotional attachment plays a negative role in predicting exhibition satisfaction, and visit frequency shows significant differences in the impact of exhibition dependence on satisfaction. The research provides specific theoretical and management enlightenment for the exhibition attachment and exhibition satisfaction of the participants. However, the model they put forward is too dependent on the subjective opinions of users and cannot objectively obtain the satisfaction of participants [5]. Lai and Wong developed a ternary reciprocal causality model to study how the knowledge created by customers affects the quality of the relationship and guarantees the perceptual performance of the report. Partial least squares analysis was performed on the data collected by 395 exhibitors in Macau. The results show that the participation of exhibitors in the preparation work has a positive impact on the three factors of relationship quality (satisfaction, trust, and commitment), and these three factors have a positive impact on the perceived performance of the report. Compared with exhibitors at public exhibitions, trust has a greater impact on exhibitors' commitments at trade shows. Research shows the integration of relationship quality theory, customer engagement theory, and social knowledge theory. Only for the three factors of the quality of exhibition participation relations, further research can be done. Yi's research draws on the theory of place attachment and puts forward an

innovative concept of exhibition attachment. Based on an in-depth literature review, a conceptual model of exhibition attachment was constructed, including push/pull motivation, attachment, and loyalty. To test the model, empirical data was collected from the participants of the trade show. The results show that there is a certain degree of correlation between the participants' push motivation and pulling motivation. Pulling motivation has a significant impact on exhibition dependence, which in turn has a positive and strong impact on the identity of the exhibition. The push-pull motives he studied have little correlation [6]. Harris and Dieringer conducted an economic analysis of a proposed zoning change that would allow the City of Treasure Island, Florida, to attract more visitors. The proposed zoning change will result in hotel renovations and new hotel construction in Treasure Island, Florida. Harris uses tax analysis to understand how this change will increase taxable property values as well as increase tourism development tax revenues [7]. Giampiccoli et al. explored the intersection between sustainable tourism and community-based tourism (CBT) and revealed a serious mismatch in the concept of sustainability and that addressing its negative environmental impacts is a fundamental requirement for tourism. A. Giampiccoli argues that sustainability in tourism should be an intrinsic and universal principle that governments should implement for all forms of tourism. In fact, traditional tourism causes most of the environmental damage and because of this, it should upgrade its role to become a solid promoter of sustainable measures that promote environmentally friendly and sustainability-friendly practices [8]. Wang et al. proposed a method to supplement fixed lockers with mobile parcel lockers to meet last-mile delivery needs. With the objective of minimizing operational costs, the location and route optimization problems of mobile delivery lockers are integrated into a nonlinear integer programming model. Embedded GA has been developed to simultaneously and optimally determine the location of delivery points, the number of mobile parcel lockers required at each delivery point, and the schedule and route of mobile parcel lockers [9]. Li et al. proposed a novel combined parametric modeling and design optimization technique for microwave components using neural networks. The proposed technique provides an iterative mechanism between artificial neural network model training and design optimization updates. This iterative mechanism is fully automated and does not require human intervention. In addition, the proposed technique overcomes the limitation of common artificial neural network optimization strategies, where the fixed training area of the artificial neural network model limits the degrees of freedom for design optimization. The proposed technique automatically expands the ANN training region until an optimized solution satisfies the user design specification [10]. In order to improve the thermal efficiency of blast furnace to achieve the goal of energy saving and consumption reduction, Zha Ng and Li optimized the blast furnace combustion control process by collecting a large amount of field data and introducing isolated point removal strategy for preprocessing with semisupervised clustering algorithm. The method has been experimentally

validated and shows significant energy savings compared with the traditional data preprocessing strategy [11]. Wang et al. proposed a multiobjective evolutionary algorithm based on a decomposition-based differential evolutionary strategy to solve the multiobjective model. A series of nondominated solutions are obtained, which represent the optimal spare parts supply solutions. In order to measure the performance of these solutions comprehensively, suitable quantitative metrics are selected and the efficiency of the obtained optimal solutions is evaluated using a cross-efficiency data envelopment analysis (DEA) model based on secondary objectives [12]. Ramazan proposed an adaptive particle filter (APF). In the APF, the adaptive fuzzy square root untraced Kalman filter (AFSRUKF) is used to generate the proposed distribution. In APF, the resampling process is done based on particle swarm optimization (PSO) in order to increase the diversity of particles. With this resampling strategy, particles with small weights are modified to particles with large weights without duplicating and eliminating particles [13].

The innovation of this paper is to use high-performance computing technology to study the dynamic mechanism of urban exhibition tourism development and establish the dynamic mechanism model of urban exhibition tourism based on the theory of tourism dynamic mechanism and tourism stakeholder theory. At the same time, it combines theoretical research and actual cases of city exhibition tourism to analyze and propose implementation plans. According to the evaluation system and the dynamic mechanism of exhibition tourism development, the dynamic characteristics of various types of power are sorted out, the development path is planned, and development suggestions are put forward. According to the weight difference of each factor of the index, the corresponding countermeasures are proposed to provide preliminary theoretical reference for tourism development researchers. From the load balancing scheduling mechanism, the overall load balancing algorithm can be divided into three categories: static load balancing algorithm, dynamic load balancing algorithm, and adaptive load balancing algorithm.

Convention and exhibition tourism is an important direction of world tourism development, an important part of urban tourism, and a model for the development of urban related industries. Convention and exhibition tourism has now become an important channel for economic and trade exchanges between countries. The convention and exhibition industry, including convention and exhibition industry, exhibition industry, and incentive tourism industry, is a huge industry in the world. Related industries with about 1% of the world's total GDP benefit from the report, and the exhibition industry contributes 8% to the global economy. Therefore, the tourism departments of various countries attach great importance to the development of exhibition tourism.

## 2. Optimization of Urban Exhibition Tourism Development

The characteristics of exhibition tourism are that exhibitors have strong spending power, the tourism industry benefits

more, tourists stay for a long time, there are more travel opportunities, and there are more potential participants, which can provide a broad platform for the development of the tourism industry.

The developmental dynamic mechanism refers to the organic collection with specific functions formed by the interaction of elements. The internal mode of interaction between elements is the structure of the system, in which each element is organized into a process for a specific purpose. The current research on the dynamic system of tourism development mainly involves the research on the dynamic mechanism of urban tourism development. These studies determine the supply of urban tourism products and the core competitiveness of urban tourism. We will jointly guide the emergence and development of urban exhibition tourism under the interaction of various dynamics such as stimulating the demand for exhibition tourism, protecting the urban environment, and supporting the country. Figure 1 is a model of the dynamic mechanism of urban exhibition tourism development.

The development of urban convention and exhibition tourism is driven by different forces, so as to realize the rapid development of convention and exhibition tourism. The factors that affect the development of urban exhibition tourism are divided into two types: driving factors (resources, economy, location, services, etc.) and supporting elements (policy, technology, culture, etc.), and their effects on urban exhibition tourism are presented as driving forces and guiding force. The figure shows the law of behavior among various factors affecting the development of exhibition tourism from the analysis of the entire city system. Among them, policy, technology, culture, economy, resources, location, and service are regarded as development driving factors.

The shift-share analysis method usually takes the economic development status of the region where the city is located as a reference, uses the gross national product as an index to evaluate the economic aggregate, and studies the change process of the regional economic industrial structure through the dynamic development process [14, 15]. When constructing the shift-share model, three indicators of the total economic volume of the case are generally selected: the first is the amount of change in the standardization industry department of the case study area based on the average growth rate of the region, that is, the deviation share  $W$ ; the second is the impact of the structural difference between the case area and the regional economy, that is, the deviation of the industrial structure from the share  $R$ ; the third is the difference in the level of economic activity caused by the difference in economic growth between the various industrial sectors within the region and the corresponding industrial sector in the case, that is, the deviation of competitiveness from the share  $L$ . Suppose that, in time  $s$ , in region  $A$ , the overall economic aggregate and industrial structure of the region have undergone some changes, and suppose that the overall economic aggregate and industrial structure of the region have undergone some changes. Assuming that the economic aggregate of this time period is  $T_{all}$  and the economic aggregate of the last year is  $F_{all}$ , then

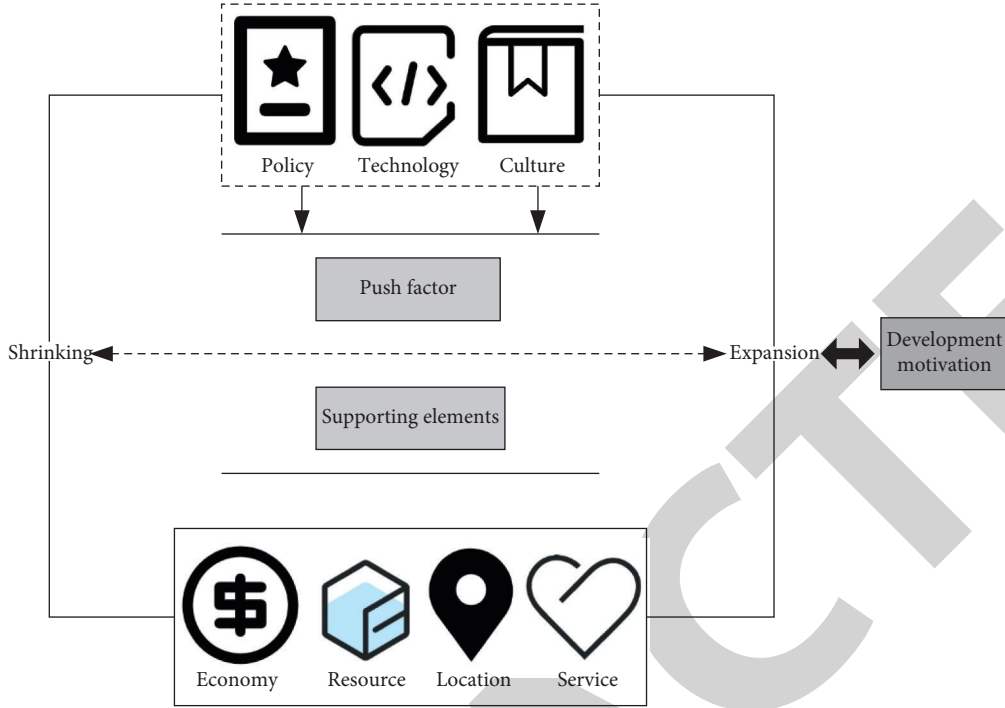


FIGURE 1: The dynamic mechanism model of city convention and exhibition tourism development.

the relevant mathematical model of the deviation-share analysis method of area  $A$  can be expressed as

$$\begin{aligned}
 H_a &= W_a + R_a + L_a = T_{all} + F_{all}, \\
 W_a &= \frac{T_a}{F_a} * T_{all} - F_{all}, \\
 R_a &= \sum_{a=1}^N \frac{T_a}{F_a} \times T_{all}, \\
 L_a &= T_{j,a} - \sum_{a=1}^n \frac{T_a}{F_a} \times (T_{all} + F_{all}).
 \end{aligned} \tag{1}$$

Among them,  $W_a$  is the deviation share of region  $a$ ,  $R_a$  is the deviation share of the industrial structure of region  $a$ , and  $L_a$  is the competitiveness share of region  $a$ .

$$\begin{aligned}
 (R + L)_a &= R_a + L_a = T_{all} - \frac{T_a}{F_a}, \\
 \eta &= \frac{T_{i,a}}{\sum_{a=1}^n (T_a/F_a) \times F_{j,a}}.
 \end{aligned} \tag{2}$$

$\eta$  is the regional competition effect index, which can more accurately analyze the status quo of the industrial structure of the economic development of region  $a$ .

High-performance computing refers to a device with a very powerful computing capability to complete a specified amount of computing work in a short period of time, mainly used to deal with complex calculation problems, and the development dynamic mechanism of exhibition tourism in this paper needs to calculate various dynamic factors in

exhibition tourism. At present, high-performance computing is widely used in scientific research fields, such as national defense and military industry, molecular chemistry, computer engineering assistance, global long-term meteorology, weather and disaster forecasting, process improvement and environmental protection, biological information technology, and medical and new drug research. Currently, the data is growing massively, from the Internet of Things to cloud computing, from big data to artificial intelligence, from Bitcoin to blockchain [16, 17], and the information revolution is rapidly promoting the development of the industry, subverting traditional concepts, and affecting around the world. Often companies need to make faster development and greater innovation in their business, accompanied by fast, stable, strong, and more demand for new computing capabilities. In the field of artificial intelligence, the rapid development of deep learning technology represented by graphics, images, audio, and video requires high-performance computing power to provide support [18, 19].

Figure 2 is the general architecture of the high-performance computing hardware platform.

With the deepening of scientific research, fierce competition, and rapid growth of data volume, digital and analog calculations are becoming more and more accurate, and the demand for computer resources is also increasing. Higher and more comprehensive requirements are put forward for data storage, computing speed, response time, and network bandwidth. Simulation has become an indispensable tool for researchers to explore the system without resorting to actual experiments. In any case, increasing the size and accuracy of the model will increase the amount of calculation [20, 21]. Consider computing, storage, and networking from three

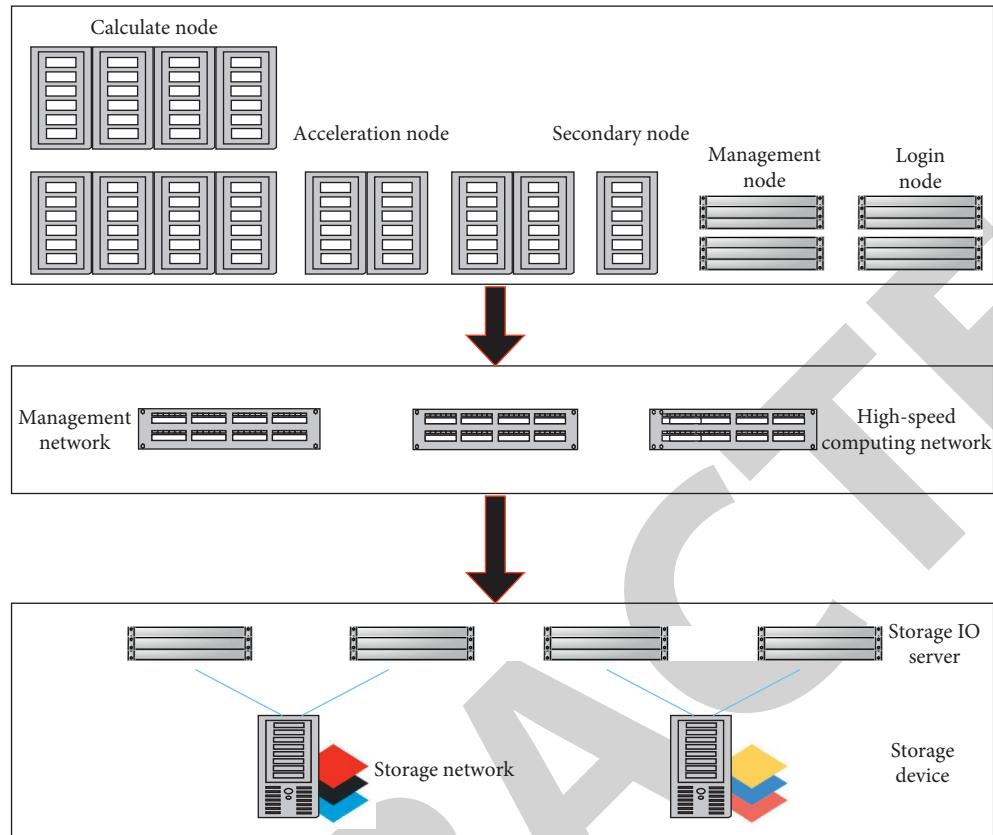


FIGURE 2: General architecture of high-performance computing hardware platform.

perspectives. IT mainly includes computer nodes, acceleration nodes, auxiliary nodes, connection nodes, and management nodes. There are three network levels in the network: high-speed computer network, storage network, and management network; separate the management levels inside and outside the area, and the management network system and business system do not share the same level [22, 23]. When the business system fails, the administrator can also manage the server through the offline management network console [24].

From computers, minicomputers, to mainframes, the rapid development of computers is still unable to meet the needs of computing. Scientific computing, network computing, terminal computing, cloud computing, supercomputing, intelligent computing, GPU (Graphics Processing Unit) computing, and other computing modes, concepts, technologies, and applications dominate the progress and development of science and technology. Quantum computing, brain-like computing, borderless computing, human-machine-object ternary fusion computing, data-intensive computing, etc. have brought computing into the multielement era [25, 26]. Calculations allow humans to explore the unknown. The future development is no longer simply a certain technology or a certain type of calculation; it must be a combination of multiple calculations, and multiple theories cooperate with each other to develop together [27]. Figure 3 shows the workflow of high-performance computing.

The terminal allocates the work and sends a message to each node to let them perform their own part of the work. The work is put into the high-performance computing unit and sent to each node at the same time, and each node is usually expected to give a result at the same time as a response. The results from each node are returned to the host application through another message provided by MPI, and then all messages are received by the application.

**2.1. Balanced Load.** Load balancing refers to the processing of a large amount of data requests or a large amount of data requests in a short period of time through resource scheduling and task allocation. It balances the distribution of load and avoids overload, so as to maximize the use of system resources, maximize system throughput, and minimize task idle time. Generally speaking, load balancing is a global optimization scheduling problem. Load balancing scheduling is classified according to types and can be divided into local and global scheduling strategies. Global scheduling strategy is divided into dynamic load balancing and static load balancing; dynamic load balancing is divided into centralized and distributed scheduling strategies, etc. Here from the load balancing scheduling mechanism, the overall load balancing algorithm can be divided into three categories: static load balancing algorithm, dynamic load balancing algorithm, and adaptive load balancing algorithm. The balanced load model is shown in Figure 4.



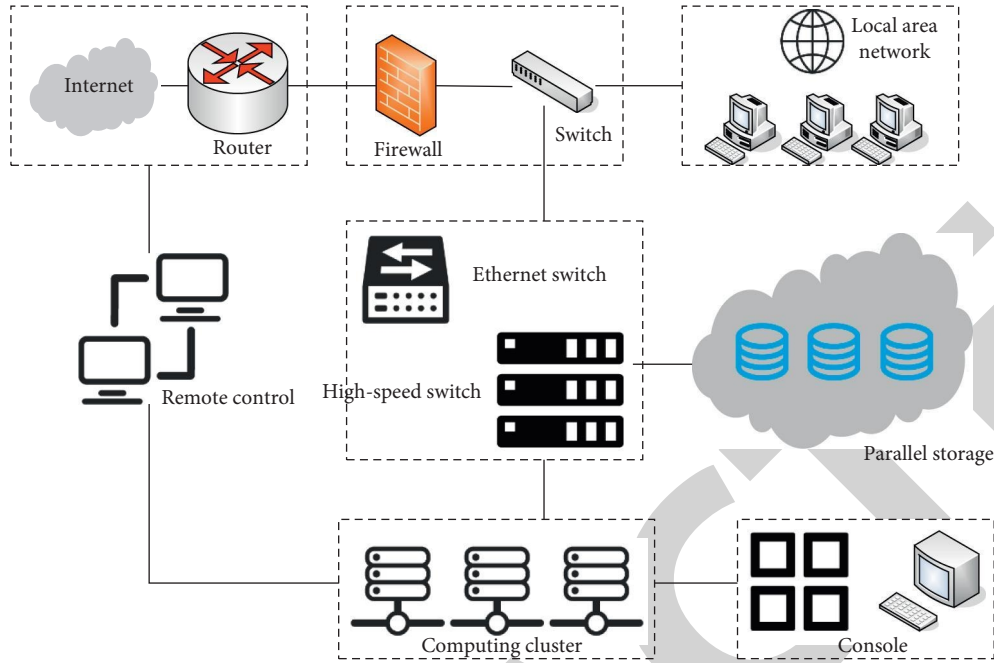


FIGURE 3: High-performance computing job workflow.

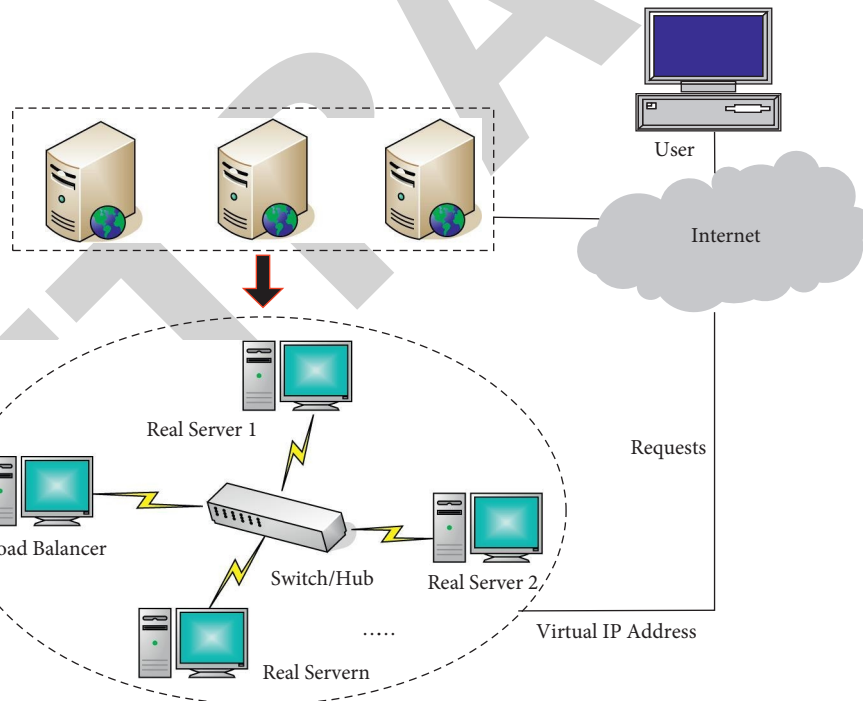


FIGURE 4: Balanced load model.

Supposing  $a$  represents an effective task scheduling strategy, then  $K(a)$  represents all the time the task set  $K$  is executed on the server node  $R$ , and the goal that the task scheduling needs to meet is to require the optimal span. Here, make  $K(a)$  as small as possible.

Let  $H$  denote the set of load balancing indicators, and set  $H = \{H_1, H_2, \dots, H_n\}$ ; then,

$$H_i(a) = \frac{1}{K(a)} \sum_{i=1}^s S_{ij} \times T_{ij}, \quad (3)$$

where  $i \in \{1, 2, \dots, n\}$ , and you can get  $0 < H_i(a) \leq 1$ , and suppose that the standard deviation  $\phi(a)$  of the load index is defined as  $\phi(a)$ ,

$$\varphi(a) = \sqrt{\frac{1}{m} \left( H_i(a) - \frac{1}{m} \sum_{i=1}^m H_i(a) \right)^2}. \quad (4)$$

According to  $0 < H_i(a) \leq 1$ ,  $0 \leq \phi(a) < 0.5$  and  $\gamma(a)$  is the standard deviation of the load index; then  $\gamma(a)$  is defined as

$$\gamma(a) = 1 - (2 \times \phi(a)). \quad (5)$$

According to  $0 \leq \phi(a) < 0.5$ ,  $0 < \lambda(a) \leq 1$  can be obtained.

According to the mathematical model, the task scheduling algorithm can be described as looking for a task scheduling strategy that makes  $K(a)$  as small as possible and  $\gamma(a)$  as large as possible, which is the optimal solution.

In order to improve performance, the solution constructed in the adaptive joint algorithm is used as the initial solution  $x$ , and then the method proposed by the new neighborhood strategy is used to improve  $x$ . First, determine and select the last scheduled task on all nodes, and the first improvement has been made. This method involves finding all the inactive time intervals of the node and trying to insert the operation of the selected task while respecting the constraints. The optimal time selection parameter represents the processing time of tasks on all nodes. The processing time from the first to the last vertex of the path is expressed as

$$WL(a, b) = \left( \sum_{i \in e_{ab}} W_{a,b} \right) + W_i. \quad (6)$$

Among them,  $W_{a,b}$  is the collection of all the data before the job; calculate the workload of each operation, and according to the total workload, sort the jobs in the order of the highest workload to the lowest workload.

$$WL(a) \begin{cases} \arg \left\{ \max_{a \in E(a)} [\delta_i^k]^j \right\}, \\ s > s_0. \end{cases} \quad (7)$$

Among them,  $s$  represents the information between paths,  $\delta_i^k$  represents the heuristic distance between them, and  $E(a)$  represents a random number uniformly distributed in  $[0, 1]$ .

$$\begin{aligned} \delta_{ij}(a+1) &= (1-g)\delta_{ij}(a) + \nabla\delta_{ij}, \quad 0 < g < 1, \\ \Delta\delta_{ij}(a) &= \sum_{k=1}^{ij} \delta_{ij}^k(a). \end{aligned} \quad (8)$$

When all the data in the job distribution process constructs a complete path from the beginning to the last vertex, each data will be marked to provide pheromone for each directional arc.

$$\nabla\delta_{ij}^k(a) = \begin{cases} WL(a), & (ij, ea) \in v^k(a), \\ 0, & \text{others.} \end{cases} \quad (9)$$

**2.2. Task Division.** Task division has a significant effect on data-intensive applications, but its ability to execute complex

data structures and control flow instructions is not good. When the problem scale is small or the frequency of use is low, different task division methods are used.

Consider the program running time; suppose  $s$  is the program running time,  $s_{\text{cpu}}$  is the CPU preprocessing time,  $s_{\text{copy}}$  is the data copy time,  $s_{\text{gpu}}$  is the kernel execution time, and  $s_{\text{ret}}$  is the return time of the abnormal data copy function. Then when the exception copy is not used, there are

$$S = S_{\text{cpu}} + S_{\text{copy}} + S_{\text{gpu}}. \quad (10)$$

When using asynchronous data copy, there are

$$s = \max \left\{ (s_{\text{cpu}} + s_{\text{ret}}), (s_{\text{copy}} + s_{\text{gpu}}) \right\}. \quad (11)$$

Neglecting the execution time of kernel and the return time of the asynchronous data copy function, it can be expressed as

$$s = \begin{cases} s_{\text{cpu}} + s_{\text{copy}}, \\ \max \{ s_{\text{cpu}}, s_{\text{copy}} \}. \end{cases} \quad (12)$$

It is not difficult to see from the formula that when the preprocessing time is close to the data copy time, the use of asynchronous data copy can shorten the program running time by half. The concurrent operation of data copy and execution has the effect of the pipeline, which can effectively hide the delay and improve the report and document classification program performance.

For a single job in the system, without considering the impact of other jobs, there may be a time interval between the occurrence of communication events during its execution. The analysis of communication characteristics in the time dimension of parallel applications mainly focuses on the interval distribution of the communication interval, and the interval is

$$\begin{aligned} g\{f(p, w+1)\} - f(p, w) &= P(a), \\ W(a) &= \{(w-1)f(w-1, a) - wf(w, a)\}, \end{aligned} \quad (13)$$

where  $g$  represents the mathematical expectation and  $P(a)$  is the proportional coefficient. If the  $a+1$  interval appears in the middle of the previous  $a$  interval, it will increase on the basis; then,

$$g\{f(1, a+1)\} - f(1, a) = \varepsilon - P(a)f(1, a) \quad (0 < \varepsilon(a+1) < 1). \quad (14)$$

It can be seen that there is an intersection of the communication intervals of multiple communication events. When there is no noncommunication interval among them, no communication event occurs.

Using the principles and steps of analytic hierarchy process, construct the quantitative evaluation index system of the urban exhibition tourism development dynamic mechanism to determine the weight of its evaluation factors. The development motivation factor is used as the first-level evaluation index, including policy, technology, culture, economy, resources, location, and service, expressed in  $T_1-T_7$ . Then the factors under the first-level indicators are



used as the second-level indicators, and the detailed evaluation indicators are shown in Tables 1 and 2.

The secondary indicators of the development momentum of exhibition tourism include economics, policies, resources, services, location, technology, culture, and other content. The level of economic and industrial development reflects the inherent support for the development of urban exhibition tourism. The urban economy has the early advantages of urban construction. Convention and exhibition tourism uses typical industries as growth engines. It is understandable that typical industrial development drives urban development. Especially as a tourist city, the development of tourism affects all aspects of urban construction. Therefore, the level of economic development and the development of typical industries directly affect the development of tourist cities. Among them, the stable and healthy development of the exhibition tourism industry requires a good market order, which requires the government to formulate relevant laws and regulations to regulate various commercial behaviors in the industry. On the basis of improving the current industry standards and regulations, the government has formulated more comprehensive qualification assessment standards, evaluation mechanisms, market competition, market supervision, and related laws, regulations, and preferential policies in the exhibition and tourism industry. It is used to coordinate the interest relationship between all parties and prevent the vicious competition in the market caused by the interests of all parties, so as to create a good market order and realize the healthy and stable development of Guilin convention and exhibition tourism.

According to the quantitative nature of the indicators, construct indicators for the dynamic development mechanism of urban exhibition tourism. The first-level indicators are elements of support and promotion, and the second-level indicators are improvements to the first-level indicators, which can more fully reflect the reporting situation of the city, judging the relative importance of each element at each level, and expressing the judgment in the form of a digital table to create a judgment matrix. The structure of the judgment matrix is shown in Figure 5.

By constructing a judgment matrix, calculating the factor weights layer by layer, and then obtaining the ranking of the importance of the factors that affect the complex problem, the evaluation of the development momentum of urban conventions and exhibitions is mainly based on the weight difference of various factors, and the corresponding development path of the convention and exhibition tourism planning will not have a major impact on development trends and changes due to factors such as time.

From the perspective of tourists' recognition, this study established a city exhibition tourism satisfaction evaluation system through questionnaire surveys and expert interviews. Through the analysis of the development factors of exhibition tourism, the development process of urban exhibition tourism is analyzed. The problems and effective ways to improve the competitiveness of city exhibition tourism are put forward. In this task, we distributed 150 questionnaires to different target groups and collected first-hand raw data

through the questionnaires. We selected evaluation indicators through on-site visits and inspections, participation in relevant exhibitions, study and research of relevant departments, and the perceived satisfaction of exhibition exhibitors, tourists, and local residents. We used the data collected from the survey to directly collect the original data. Figure 6 shows a framework for analyzing and investigating MICE tourism satisfaction.

### 3. Optimization Results of Urban Exhibition Tourism Development

The exhibition tourism industry is a huge driving force for urban development and product brand building. With the increasing frequency of economic and trade cooperation and exchanges in various regions, large-scale exhibitions in various cities have become important promoters of economic and trade cooperation, economic introduction, and products going abroad. Convention and exhibition tourism has gradually developed into an important industry in the economic development of each city. At the same time, according to the increasingly fierce competition in urban development location and the continuous segmentation of products and markets, it has a clear position in the development of exhibition tourism. The positioning of these cities for the development of exhibition tourism undoubtedly determines the trend and positioning of the city in the future development. It also determines the trend pattern of future exhibition tourism development and the overall strength of the city's development. The tourism development of city A in recent years is shown in Figure 7.

Figure 7 shows the tourist arrivals and growth rate of city A in recent years. It can be seen that, from 2015 to 2019, the annual tourist arrivals and tourism growth rate of city A have increased significantly. According to statistics, the annual profit of exhibition tourism in the tourism market of city A is more than 20%–25%, and there are more than 400 large-scale exhibitions each year, which is in the upper reaches of the country.

From Figure 8, combined with the number of tourists in Figure 7, it can be seen that the exhibition tourism market in city A has shown great development potential. Among tourists, conference tourists account for 10.3% of the total tourists, second only to vacation and sightseeing tours. Its unique tourism resources, rich leisure, and entertainment activities and perfect service system have exerted a strong pull on the development of exhibition tourism in city A.

According to the first- and second-level evaluation indicators of the development momentum of exhibition tourism constructed in the previous article, the weight coefficients contained in each indicator are calculated. Table 3 is the judgment matrix of the first-level index weight coefficient, which determines the weight of each evaluation factor and then improves and optimizes its development motivation mechanism, so that the exhibition tourism can develop in a healthy and orderly manner.

Analyzing the weights of the first-level evaluation indicators, it can be seen that tourism resources and economic factors and policy factors have a relatively high weight.

TABLE 1: Evaluation indexes for the development motivation mechanism of convention and exhibition tourism.

Factor	First-level evaluation index	Factor	Secondary evaluation index	Sort
$T_1$	Economy	$S_1$	Capital investment	3
		$S_2$	Featured industry cluster	
		$S_3$	International level	
$T_2$	Resource	$S_4$	Travel resources	1
		$S_5$	Exhibition facilities	
		$S_6$	Infrastructure	
		$S_7$	Human resources	
$T_3$	Location	$S_8$	Political culture	2
		$S_9$	Natural economy	
		$S_{10}$	Regional market	
$T_4$	Service	$S_{11}$	Travel services	6
		$S_{12}$	Expo services	

TABLE 2: Construction of evaluation indexes for motive mechanism of exhibition tourism development.

Factor	First-level evaluation index	Factor	Secondary evaluation index	Sort
$T_5$	Policy	$S_{13}$	Industrial policy	4
		$S_{14}$	Exhibition system	
		$S_{15}$	Government resources	
$T_6$	Culture	$S_{16}$	Citizen quality	7
		$S_{17}$	The image of a city	
		$S_{18}$	City safety	
$T_7$	Technology	$S_{19}$	Exhibition technology	5
		$S_{20}$	Management technology	
		$S_{21}$	Information technology	

( $T$ )	$S_1$	$S_2$	$S_3$	...	$S_n$
$S_1$	$S_{11}$	$S_{12}$	$S_{13}$	...	$S_{1n}$
$S_2$	$S_{21}$	$S_{22}$	$S_{23}$	...	$S_{2n}$
...	...	...	...	...	...
$S_n$	$S_{n1}$	$S_{n2}$	$S_{n3}$	...	$S_{nn}$

FIGURE 5: Evaluation index judgment matrix.

Among them, the weight coefficient of tourism resources is 0.2411, which has the highest weight. After the judgment and analysis of the index matrix, and the consistency test of the single-level and total rankings, the weight coefficients of the secondary indicators are analyzed. Figure 9 shows the weights of the secondary evaluation indicators of the exhibition tourism development dynamic mechanism.

From the perspective of system science, the analysis of tourism power mainly includes the attraction system based on tourism demand, the attraction system based on tourism

supply power, the tourism promotion power system, and the supporting power formed by its development. Therefore, the development of exhibition tourism has common characteristics. Secondly, from the internal analysis of the dynamic system, these factors are not independent, nor are they purely linear. There is a certain causal effect and feedback influence between them. Refer to the previous evaluation index composition, combined with the relevant data of the current situation, to obtain the power index of the power factor index of the exhibition tourism development.

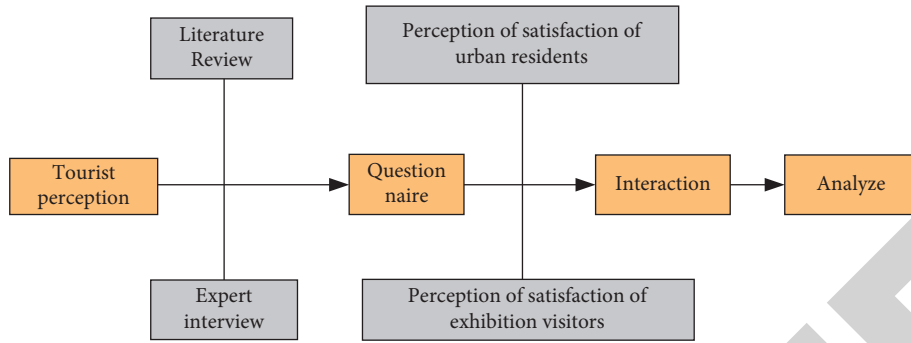


FIGURE 6: The analysis and research framework of exhibition tourism satisfaction.

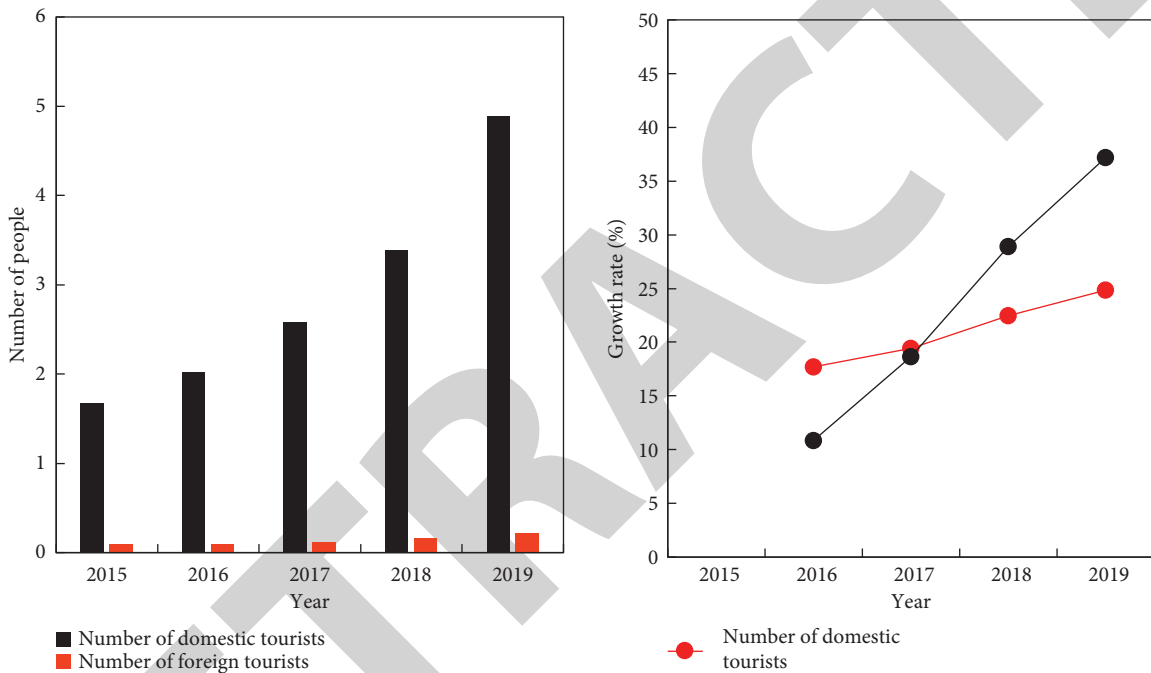


FIGURE 7: Number of domestic and foreign tourists in Hangzhou from 2015 to 2019.

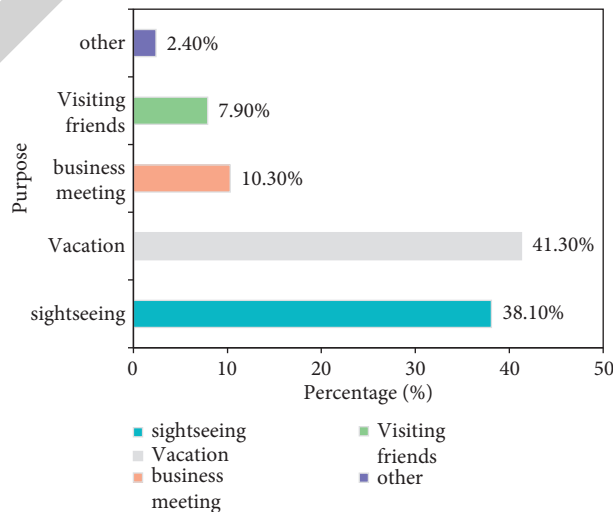


FIGURE 8: x city's tourist purpose in recent years.

TABLE 3: Judgment matrix of first-level index weight coefficient.

Power	$T_1$	$T_2$	$T_3$	$T_4$	$T_5$	$T_6$	$T_7$	$W$
$T_1$	1.97	1.85	2.68	2.12	1.11	2.08	2.86	0.1523
$T_2$	1.55	2.85	1.98	1.82	2.96	2.98	1	0.2411
$T_3$	1.66	1.33	2.2	1.11	2.1	2.05	1.36	0.1804
$T_4$	2.79	2.39	2.58	2	1.42	1.33	2.59	0.0993
$T_5$	2.89	1.76	2.5	1.17	2.33	1.8	1.73	0.1458
$T_6$	2.6	1.66	2.02	1.01	1.53	1.6	2.05	0.0769
$T_7$	1.32	2.75	1.71	1.54	1.23	2.97	2.82	0.1042

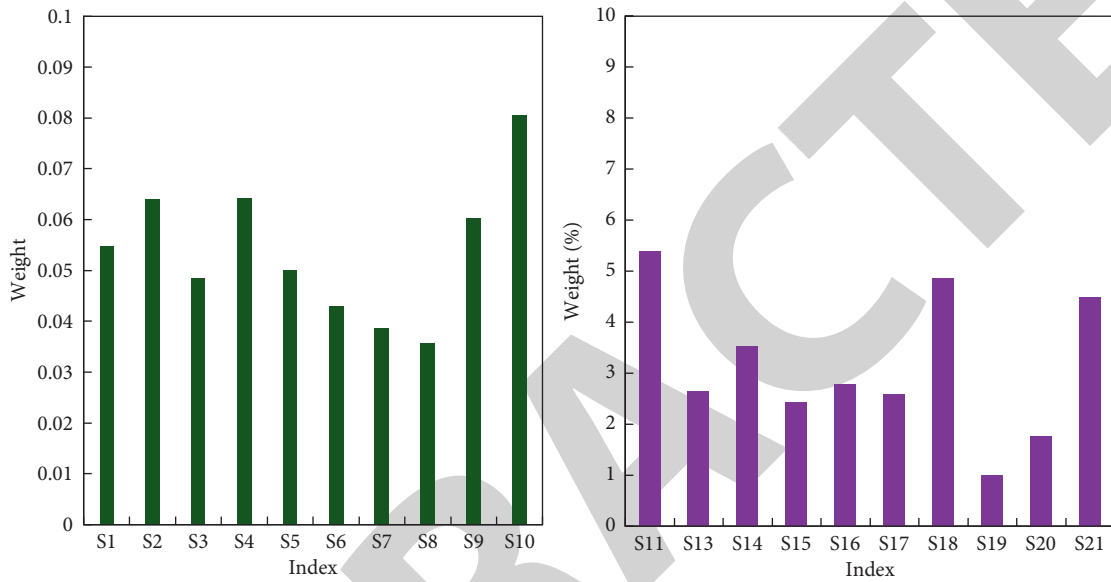


FIGURE 9: The weight of the secondary evaluation index of the development dynamic mechanism of exhibition tourism.

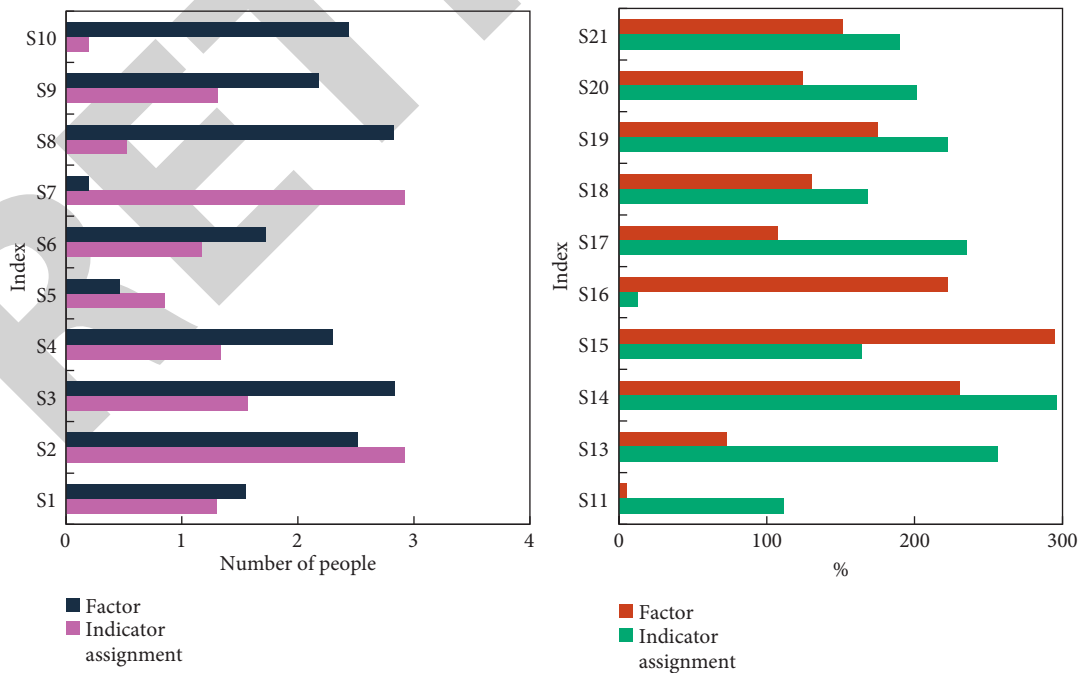


FIGURE 10: Motivation factors and index assignments for the development of exhibition tourism.

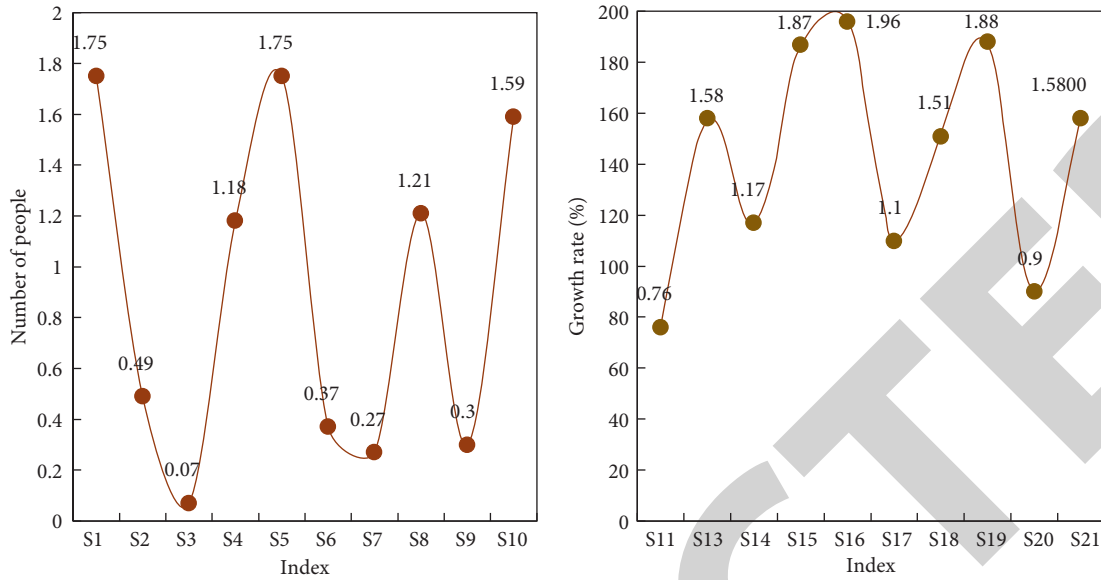


FIGURE 11: Development momentum index.

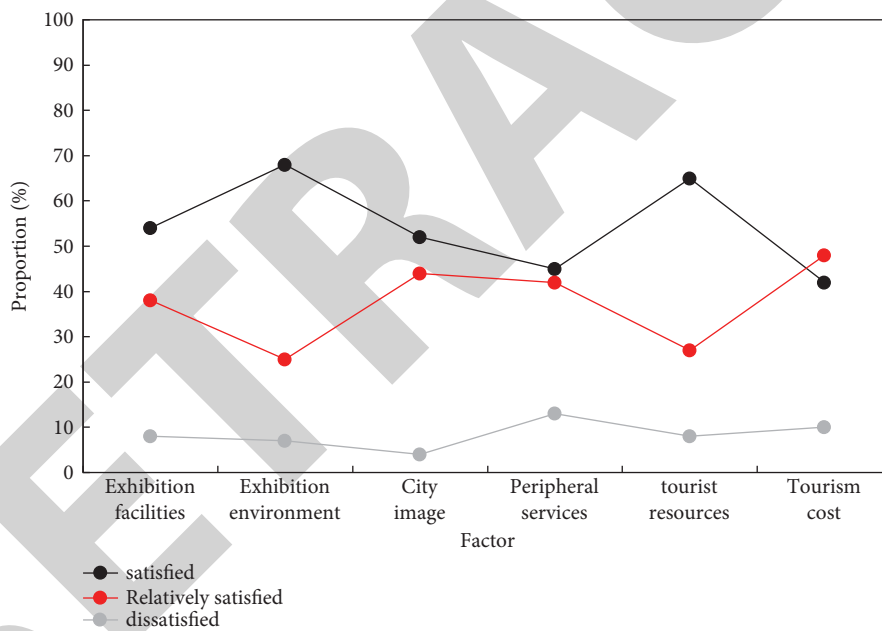


FIGURE 12: Satisfaction statistics of MICE tourism.

The power index and the index assignment are shown in Figures 10 and 11.

Use index assignment to determine the weight of dynamic factors and determine the degree of influence of different dynamic factors on the development of exhibition tourism. That is, the evaluation of the dynamic factor and the index assignment mainly adopt the analytic hierarchy process.

After calculating the weights of the indicators, the indexes of the dynamic factors at various levels are obtained. It can be seen that although the tourism resources of city A are relatively rich, the actual power is more, and the location and environmental factors are relatively strong. However, in

view of the current driving force of the development of exhibition tourism in which resources play a major role, the driving force of resource factors on the development of exhibition tourism is still relatively strong.

The most important thing in MICE tourism is the degree of satisfaction of tourists during their travel. Only when the tourists are satisfied can they attract more tourists to travel. Through the survey of tourist satisfaction in exhibition tourism, we can analyze and study the development direction of Hu Zhan tourism and the improvement of exhibition tourism service quality. Figure 12 is the questionnaire analysis on the satisfaction of exhibition tourism.

In the questionnaire for residents' satisfaction, urban factors, exhibition facilities, exhibition environment, tourism resources, tourism costs, and other indicators are set to analyze the satisfaction of exhibition tourism. Through the analysis of survey statistics, we can use evaluation indicators to make it clearer that customers attach great importance to those issues in the development of exhibition tourism, but the customer perception evaluation is very low. These need to be strengthened and improved by the exhibition management department and related organizers.

#### 4. Conclusions

This paper analyzes the dynamic mechanism of urban exhibition tourism development through high-performance computing and conducts research on the optimization countermeasures of exhibition tourism. Under the premise of clarifying the differences in the development of city convention and exhibition tourism, the dynamic factors of the development of convention and exhibition tourism industry are analyzed. Regarding the main characteristics and content of the seven main driving factors of exhibition tourism development, refer to the relevant evaluation index system to subdivide the driving factors of exhibition tourism development, economic factors, selection policy factors and resource factors, and other location factors. Taking into account the factors of the investigation, the concept and composition of each element are refined layer by layer, and finally a multifactor evaluation system for the development of exhibition tourism is formed. We used surveys and detailed layering procedures to determine the weights of all levels of measurement. From the weighted results, we can see the characteristics of the power structure of the city's exhibition tourism in the emergence stage. The load balancing algorithm is mentioned in the article, and load balancing is a global optimization scheduling problem. Combined with the principle of dynamic factors, it analyzes and summarizes the characteristics and main development paths of exhibition tourism under the combined effect of various factors. However, there are still some shortcomings in the article research, such as the lack of depth in the current research on the dynamic mechanism of the development of exhibition tourism in the city, and the selection of evaluation factors and mechanism research on the dynamic mechanism of exhibition tourism are not perfect. We hope it can be improved in the future work.

#### Data Availability

The data used to support the findings of this study are available from the author upon request.

#### Conflicts of Interest

The author declares no conflicts of interest.

#### Acknowledgments

This research study was sponsored by Guangdong Social Science Planning Project, Research on Coexistence and

Cooperation in the Context of Guangdong-Hongkong-Macau Greater Bay Area (no. GD18XGL29), and Characteristic Innovation Project of Universities in Guangdong Province in Year 2020, Research on Resilience Assessment and Innovative Development of Tourism Enterprises in Guangdong-HongKong-Macau after Public Health Event (no. 2020WTSCX161).

#### References

- [1] T. Bridi, A. Bartolini, and M. Lombardi, "A constraint programming scheduler for heterogeneous high-performance computing machines," *IEEE Transactions on Parallel and Distributed Systems*, vol. 27, no. 10, pp. 2781–2794, 2016.
- [2] N. Jagiella, D. Rickert, F. J. Theis, and J. Hasenauer, "Parallelization and high-performance computing enables automated statistical inference of multi-scale models," *Cell Systems*, vol. 4, no. 2, pp. 194–206, 2017.
- [3] K. Wang, A. Kulkarni, M. Lang, D. Arnold, and I. Raicu, "Exploring the design tradeoffs for extreme-scale high-performance computing system software," *IEEE Transactions on Parallel and Distributed Systems*, vol. 27, no. 4, pp. 1070–1084, 2016.
- [4] J. Mullen, C. Byun, V. Gadepally, and S. Samsi, "Learning by doing, high performance computing education in the MOOC era," *Journal of Parallel and Distributed Computing*, vol. 105, pp. 105–115, 2017.
- [5] X. Fu, X. Yi, F. Okumus, and W. Jin, "Linking the internal mechanism of exhibition attachment to exhibition satisfaction: a comparison of first-time and repeat attendees," *Tourism Management*, vol. 72, pp. 92–104, 2019.
- [6] I. Lai and J. Wong, "From exhibitor engagement readiness to perceived exhibition performance via relationship quality," *Journal of Hospitality and Tourism Management*, vol. 46, pp. 144–152, 2020.
- [7] R. L. Harris and S. Dieringer, "The economic impact of rezoning to increase tourism," *Tourism*, vol. 65, no. 1, pp. 103–114, 2017.
- [8] A. Giampiccoli, O. Mtapuri, and A. Druzewska, "Investigating the intersection between sustainable tourism and community-based tourism," *Tourism*, vol. 68, no. 4, pp. 415–433, 2020.
- [9] Y. Wang, M. Bi, and Y. Chen, "A scheduling strategy of mobile parcel lockers for the last mile delivery problem," *Promet - Traffic - Traffico*, vol. 32, no. 6, pp. 875–885, 2020.
- [10] Y. Li, M. Xu, X. Wen, and D. Guo, "The role of internet search index for tourist volume prediction based on GDFM model," *Tehnicki vjesnik-Technical Gazette*, vol. 27, no. 2, pp. 576–582, 2020.
- [11] C. N. Zha Ng and Y. R. Li, "Optimization analysis based on intelligent control of the process of the blast furnace," *Metallurgija -Sisak then Zagreb-*, vol. 58, no. 1, pp. 7–10, 2019.
- [12] Y. Wang, Q. Shi, Z. You, and Q. Hu, "Integration methodology of spare parts supply network optimization and decision-making," *Promet - Traffic - Traffico*, vol. 32, no. 5, pp. 679–689, 2020.
- [13] H. Ramazan, "An adaptive particle filter based on PSO and fuzzy inference system for nonlinear state systems," *Automatika*, vol. 59, no. 1, pp. 94–103, 2018.
- [14] R. Guerra, E. Martel, J. Khan, S. Lopez, P. Athanas, and R. Sarmie, "On the evaluation of different high-performance computing platforms for hyperspectral imaging: an OpenCL-based approach," *Ieee Journal of Selected Topics in Applied*



## Research Article

# Research on Subsidy Strategy of Shared Accommodation Platform under the Background of Big Data Based on Evolutionary Game

Meng Xiao  and Chang Zhai

*School of Management, Shenyang University of Technology, Shenyang 110021, China*

Correspondence should be addressed to Meng Xiao; [mengxiao@sut.edu.cn](mailto:mengxiao@sut.edu.cn)

Received 31 January 2022; Revised 27 February 2022; Accepted 28 February 2022; Published 1 April 2022

Academic Editor: Daqing Gong

Copyright © 2022 Meng Xiao and Chang Zhai. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

To attract customers and landlords to join the platform is an important problem to be solved when expanding the size of the shared accommodation platform, and subsidy policies have been proven to be an effective approach in many areas of sharing. In order to analyze the strategies of all parties in the subsidy policy of the shared accommodation platform, a three-party evolutionary game model of shared accommodation platform, consumers, and landlords was built. The strategy stability of each game subject was analyzed, and the equilibrium point stability was explored based on Lyapunov's method. The game model and influencing factors were simulated and analyzed by MATLAB 2016. Results demonstrated that the consumers and landlords were promoted to join the shared accommodation platform by increasing subsidies to consumers and landlords and reducing consumers' time costs and landlords' service fees from the early stage to the rapid development stage of the shared accommodation platform; with the increase of the proportion of landlord's share in expenses and the decrease of the opportunity loss of platform nonsubsidizing, the strategy of the platform gradually evolves from subsidy to nonsubsidy. The conclusions of this study provide guidance for the mature path of the shared accommodation platform and also give suggestions for the development of the sharing economy.

## 1. Introduction

The comprehensive application of emerging technologies such as the Internet and big data in the market brings special operation mode, effective realization tools, and high-quality value creation, which have opened up a new road for the development of market industries and accelerated the evolution of the global economic format. In recent years, benefiting from technological advances and the increasing awareness of people's awareness towards sustainability, the ranges of shared products have increased dramatically [1]. Sharing has turned from a "nonreciprocal prosocial behavior" [2] to a unique economic model [3], which brings innovation and efficiency to local development [4]. According to Juniper Research, the global market for the sharing economy is expected to reach \$40.2 billion in 2022 [5]. Sharing platforms are springing up all over the world, and the emerging business model of "Internet Plus" is

gradually becoming the important engine for global economic upgrading and transformation. The sharing economy has started to develop in many industries around the world, such as Airbnb, an online travel and housing rental company established in 2008, and Uber, an online car service established in 2009. As an important trend for future development [6], the sharing economy has had a profound impact on the social economy to change people's consumption behavior and consumption patterns [7] and develop many new models by integrating with various industries [8]. Among the industries, the development of the shared accommodation industry has been remarkable [9].

The traditional accommodation industry is an important sector of the service industry, providing entertainment and dining services to travelers. However, with rapidly changing market conditions and competition, shared accommodation is a new way for the industry to develop as a result of the deep integration of the accommodation industry with

information technology. According to the “China Sharing Economy Development Report 2020,” the market size of shared accommodation in China was about 22.5 billion yuan, increasing of 36.4% over the previous year, and the revenue of the shared accommodation industry accounted for 7.3% of the total revenue of the accommodation industry in 2019. The growing shared accommodation industry is bound to have a profound impact on the development of the industry. Therefore, it is important to analyze the development law of shared accommodation and understand the model in the development stage, which will be guiding significance for promoting the shared accommodation industry and socioeconomic development.

As a disruptive innovation, shared accommodation has become an important model to allocate urban accommodation resources, revitalize the rural tourism industry, and activate the development potential of the service industry [10–12]. Some companies have already achieved success in shared accommodation, represented by HomeAway and Airbnb Inc. [13]. The research on shared accommodation has gradually increased since 2013, with the market size of shared accommodation expanding [14]. The development of the shared accommodation industry in many areas imitated Airbnb, but the process is difficult for the trust barriers and legal regulation problems among various stakeholders [15]. Some scholars have proposed that the development of shared accommodation in China draws on the successful experience of Airbnb [16, 17], but it also faces the dilemma of “acclimatization” [18]. Therefore, how to keep the market alive of shared accommodation has become a problem that needs to be solved by shared accommodation companies in the global.

Shared accommodation is a type of nonstandard accommodation in which homeowners rent out their idle rooms to travelers for a short period through an online network platform [19]. Compared to traditional accommodation, shared accommodation has a wider variety of room types and caters to travelers’ desire to “live like a local” with an authentic destination experience [20], which makes it popular among customers [21]. There are three parties involved in shared accommodation: the landlord (the operator providing the room for rent), the shared accommodation platform, and the consumer (the renter) [22]. The shared accommodation platform is the intermediary connecting landlords and consumers, increasing the efficiency for matching and reducing the search costs. As the intermediary organization, the platform charges the service fee for both landlords and consumers. Due to the characteristics of nonneutral, cross-network externalities in bilateral market pricing, shared accommodation platform usually charges the service fee on one side and no fees on the other [23]. However, compared with hotel staff, most of the landlords attracted by shared accommodation platforms are untrained, nonprofessional training staff [24], and they also face ordinary consumers. There are many problems such as short development time, weak foundation, small development size, low participation enthusiasm in the development of the shared accommodation platform. Some scholars’ studies have found that the profit has a greater influence on

professional landlords in the operation process and the price level of their rooms is higher by comparing the operating conditions of the sellers operating a single room with those operating multiple rooms at the same time [25, 26]. There is a lack of research on how to change landlords’ pricing strategies and develop reasonable pricing strategies.

As a successful sharing economy model, the shared accommodation platform is widely recognized in the world, and the reasonable operation strategy is the key to the development of the platform. Given the characteristics of sharing platforms, the products served by the sharing economy are different from traditional bilateral markets. In order to expand the size of the platform, many platforms consider using subsidies to regulate the balance between the supply and demand of bilateral users during their development phase, which has laid a good foundation for the successful operation of platforms, for example, Uber subsidized drivers and DiDi subsidized passengers. Subsidies can also be used as an incentive method to promote the increase of the number of participants in the development of shared accommodation platforms. It has been studied in the field of sharing economy platforms that the equilibrium market size obtained by subsidizing suppliers is larger than that obtained by subsidized customers [27]. As shared accommodation platforms adopt subsidies to expand the market size, subsidy policies need to be updated accordingly. Therefore, it is relevant and urgent to research the subsidy policies of shared accommodation platforms.

In order to promote the development of the accommodation industry in the context of big data and build a bilateral accommodation platform based on information technology, to make up for the insufficiency of previous research with two parties as game subjects, the behavioral strategies of all parties in the operation of the shared accommodation platform should be more comprehensive and systematic considered. This paper adopts evolutionary game theory to analyze the evolutionary mechanism and process among platform, landlord, and consumer under subsidy policy of the platform from a dynamic perspective and explores the influences of subsidy amounts, landlord service fee, consumer time cost, and the proportion of landlord’s share in expenses on system stability. Finally, corresponding countermeasures are suggested for the subsidy mechanism of the platform. This paper answers the following three key questions: (1) How to determine the strategic conditions for shared accommodation platforms to conduct subsidies and the utilities obtained by consumers and landlords participating in shared accommodation. (2) How to achieve the equilibrium strategies of shared accommodation platforms, landlords, and consumers under platform subsidy mechanism. (3) To explore the factors affecting the equilibrium strategies.

## 2. Related Literature

Most of the research on shared accommodation focuses on the factors of willingness to participate of consumers and landlords and the factors influencing the price of shared accommodation.

Many scholars have explored consumers' willingness to participate based on different perspectives, and the studies have found that the material, energy-saving, psychological, and social interaction willingness promote consumers to participate in shared accommodation [28]. Extrinsic material benefits positively attract consumers to participate in shared accommodation platforms [29], and low-priced housing is the choice of most frugal consumers, which explains the attitude of consumers to use Airbnb [30]. Consumers tend to use rather than own products, which avoids the waste of social resources, and such behavior is expected to alleviate social problems such as excessive consumption and environmental pollution by reducing the cost of economic coordination [31]. Trust, familiarity, loyalty, authenticity, and hedonistic can also influence consumer participation in shared accommodation [32–34]. Social interactivity attracts consumers to share and helps to develop new social relationships during consumption [29, 35], enhancing social interaction stickiness [36]. Although the consumers' willingness to participate is affected by many factors, few studies consider the cross-network externalities between consumers and landlords.

Existing studies on landlords' willingness to participate are mostly centered on the economic value, functional value, security value, and trust value dimensions. The prerequisite for landlords to join the platform is the platform's sustainable development prospects, functional values of operational compatibility, and ability to respond to crises and solve problems [37]. By registering room resources online, the rental of idle rooms can be realized, and the participation enthusiasm of landlords can be stimulated by monetary value so that the landlords can directly obtain economic benefits [38, 39]. But landlords still have doubts about the safety and security of personal property in the platform [40]. Trust is the key factor to overcome uncertainty and reduce [41], so the importance of information security governance is particularly prominent [42]. It is necessary to clarify the influencing factors of landlords' participation in home-sharing so that the platform can formulate corresponding strategies to attract landlords to join.

The factors influencing the pricing of shared accommodation platforms are mainly the structure characteristics and accommodation facility features of the room, location characteristics, trust, and sociality [43–45]. The rental price is mainly determined by the landlord [43]. At the same time, the rental price is affected by the comment score and room quality, with higher pricing for high star rating listings [46]. The improvement in one room quality has the spillover effect on the prices of the other room [47]. There is a correlation between external factors and room rental prices. The internal facilities of the room and the supply of parking spaces have a positive impact on the room rental price, while instant booking has a negative effect on room rental prices [43, 48]. The authenticity of the photo has a significant impact on the rental price of the shared room; the more authentic the photo, the more purchase opportunities and economic benefits it brings [49]. However, few studies have discussed how to formulate reasonable strategies for sharing accommodation platforms to take into account the interests of all

parties in sharing accommodation and achieve the optimal system utility.

The existing research methods on shared accommodation platforms are mainly to analyze the dynamic optimal allocation of idle rooms in the platform [50] and the game research between the participants of shared accommodation. However, most of the game models are studied from the perspective of two subjects, and there are few comprehensive studies on shared accommodation platforms, consumers, and landlords. The dynamic evolution law of the three stakeholders in the development process of the shared accommodation platform has not been deeply discussed; especially, the subsidy strategy of shared accommodation platforms has not been well explained. This study attempts to fill the gap by analyzing the impact of subsidy policies of shared accommodation platforms on the dynamic evolution of strategies among platform, consumer, and landlord, explores the evolution stage and influencing factors of platform subsidy implementation, and puts forward corresponding suggestions for platform development.

### 3. Materials and Methods

*3.1. Problem Description.* The shared accommodation platform can realize the sharing of information and resources and promote the rapid development of the shared accommodation industry. Currently, the shared accommodation industry is in the development stage, the platform's capacity is insufficient, the landlords are not dependent on the platform, and the consumers have limited access to information about room resources. Therefore, shared accommodation platform companies can give some subsidies for landlords and consumers to promote them to join the platform. The landlords share their idle rooms through the platform and obtain certain utilities, and consumers find satisfactory rooms through the shared accommodation platform. In this study, "platform" refers to the shared accommodation platform, "consumer" refers to the demander of the rental idle room, and "landlord" refers to the supplier of the rental idle room. The landlord ( $L$ ) posts his or her idle room on the shared accommodation platform ( $P$ ), and the consumer ( $B$ ) rents it through the platform. The platform charges fees to the landlord and the consumer.

The shared accommodation platform, landlords, and consumers are very important stakeholders in the development of the shared accommodation industry, and the three parties are connected through the rooms for rent, as shown in Figure 1.

*3.2. Model Assumptions.* In order to build the subsidy strategy model of a shared accommodation platform based on the evolutionary game in the context of big data, each party's strategy and the equilibrium point of the evolutionary game system are analyzed, and the following assumptions are made.

*Assumption 1.* In the network transaction system consisting of "shared accommodation platform-consumer-landlord," each game participant is finite rational, and their

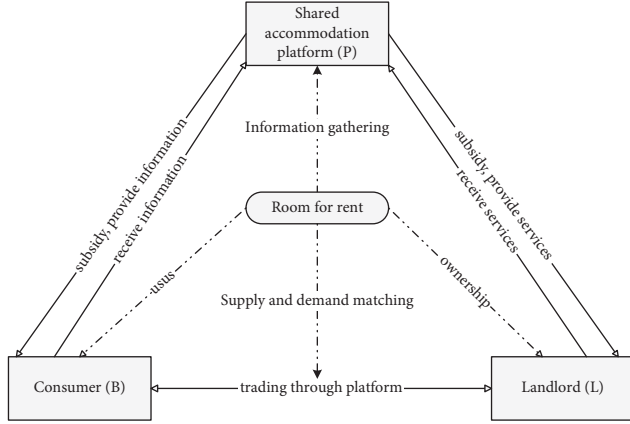


FIGURE 1: Relationship among shared accommodation platform, landlord, consumer, and room for rent.

information is asymmetric. In the game, the behaviors of the participants are random and interactional, which means that the participants cannot obtain the optimal strategies through one game but needs to go through trial and error and learn and improve their past strategies in the multiple rounds game. This means that each participant cannot obtain the optimal strategy through one game and needs to improve their strategies through continuous trial and error and learning in multiple rounds of games so as to make behavioral decisions that best match the current situation.

*Assumption 2.* In this paper, the goal of all parties is to maximize benefits under the subsidy policy of the platform. The platform determines whether to provide subsidies to consumers and landlords based on profit maximization; consumers decide whether to participate in sharing platform transactions based on their own needs to obtain the maximum utility of participating in sharing accommodation; landlords choose whether to join the platform based on the maximum net income of participating in the sharing platform. The maximum amounts of subsidy provided by the platform to landlords and consumers are, respectively,  $S_L$  and  $S_B$ . The platform determines the intensity of subsidies according to the participation of landlords and consumers, and the proportion of subsidies provided by the platform to consumers and landlords is, respectively,  $\alpha$  and  $\beta$ . Subsidies are given to landlords and consumers in the form of coupons or cash.

*Assumption 3.* In the process of operation, the sharing platform will charge corresponding fees to the landlords and consumers.  $P_B$  is the total cost paid by consumers to participate in shared accommodation, including the service fee charged by the platform and the fee charged by the landlord.

The proportion of landlord's share in expenses paid by consumers is  $\phi$ . The fee charged by the landlord is  $\phi P_B$ , and the service fee charged by the platform is  $P_B(1 - \phi)$ .  $P_L$  is the service fee charged by the platform to the landlord, and the service includes priority recommendation of room information, taking photos of the room, and providing smart locks.

*Assumption 4.* The strategy choice of shared accommodation platform is subsidy and nonsubsidy, the proportion of platform implementing subsidy is represented as  $x(0 \leq x \leq 1)$ , and the proportion of implementing nonsubsidy is denoted as  $(1 - x)$ ; the strategy choices of consumer are participation and nonparticipation, the proportion of consumer implementing participation is represented as  $y(0 \leq y \leq 1)$ , and the proportion of implementing nonparticipation is denoted as  $(1 - y)$ ; the strategy choices of the landlord are sharing and unsharing, the proportion of landlord implementing sharing is represented as  $z(0 \leq z \leq 1)$ , and the proportion of implementing unsharing is denoted as  $(1 - z)$ , where  $x, y, z \in [0, 1]$ .

*Assumption 5.* The purpose of adopting the subsidy is to enable consumers and landlords to participate in the shared accommodation platform, and the platform mainly obtains benefits by charging landlord service fees and the proportion of consumers paying room fees. Therefore, the initial subsidy of the platform needs to be more than the reputation benefit in order to better attract consumers and landlords to join the shared accommodation, that is,  $\alpha S_B > R_B$  and  $\beta S_L > R_L$ . Due to more contact with landlords on the sharing platform, the cross-network externalities utility of consumers received from participating in shared accommodation is more than the price paid by consumers, that is,  $N_B > P_B$ .

*3.3. Model Building.* The parameters of this paper are described in Table 1.

According to the above assumptions, the shared accommodation platform, consumer, and landlord are the three main players in the game, and the game model benefit matrix is shown in Table 2.

### 3.4. Model Analysis

*3.4.1. Stability Analysis of Shared Accommodation Platform Strategies.* The expected utility of shared platform selecting subsidy is  $\pi_x$ , the expected utility of shared platform selecting nonsubsidy is  $\pi_{1-x}$ , then the average expected utility of the shared platform is  $\bar{\pi}_x$ , and  $\bar{\pi}_x = x\pi_x + (1 - x)\pi_{1-x}$ , which are defined as follows:

$$\begin{aligned} \pi_x &= yz * [-C_p + P_B(1 - \phi) + P_L + R_B + R_L - (\alpha S_B + \beta S_L)] + (1 - y)(1 - z) * (-C_p) + (1 - y)z \\ &\quad * (-C_p + R_L + P_L - \beta S_L) + y(1 - z) * (-C_p + R_B - \alpha S_B). \\ \pi_{1-x} &= yz * [-C_p + P_B(1 - \phi) + P_L - H] + (1 - y)(1 - z) * (-C_p - H) + (1 - y)z * (-C_p + P_L - H) + y(1 - z) * (-C_p - H). \end{aligned} \quad (1)$$



TABLE 1: Parameter description.

Parameter	Description
$C_P$	Platform operating costs
$R_B$	Reputational benefits of platform subsidizing to consumers
$R_L$	Reputational benefits of platform subsidizing to landlords
$S_B$	Maximum subsidy amount of the platform to consumers
$A$	The proportion of subsidies provided by the platform to consumers, $\alpha \in [0, 1]$
$S_L$	Maximum subsidy amount of the platform to landlords
$B$	The proportion of subsidies provided by the platform to landlords, $\beta \in [0, 1]$
$H$	Opportunity loss of platform nonsubsidizing
$P_L$	Fees charged to landlords for the platform services
$P_B$	Fees paid by consumers participating in shared accommodation platforms
$\Phi$	The proportion of landlord's share in expenses paid by consumers, $\emptyset \in [0, 1]$
$N_B$	Cross-network externalization utility gained by consumers with landlords on the shared platform
$C_B$	Time cost of consumers
$B$	Fees paid by consumers for renting room through offline channels
$U_B$	The utility of platform service for the consumer
$N_L$	Cross-network externalities utility gained by landlords with consumers on shared platform
$U_L$	The utility of platform service for the landlord
$C_L$	The cost of room and service provided by the landlord
$A$	Income by the landlord with unsharing room in the platform

According to the Malthusian dynamic equation, the replication dynamic equation of shared accommodation platform is obtained:

$$F(x) = \frac{dx}{dt} = x(1-x) * [y * (R_B - \alpha S_B) + z * (R_L - \beta S_L) + H],$$

$$F'_x(x) = \frac{\partial F(x)}{\partial x} = (1-2x) * [y * (R_B - \alpha S_B) + z * (R_L - \beta S_L) + H].$$

According to the stability theorem of the replication dynamic equation, the evolutionary stability strategy of the shared platform must be satisfied with the conditions that  $F(x) = 0$  and  $F'_x(x) < 0$ . And the threshold of shared platform strategy transformation is  $y^* = -((z * (R_L - \beta S_L) + H) / (R_B - \alpha S_B))$ .

**Proposition 1.** *In the evolution process, when  $0 < y < y^* < 1$ , the stabilization strategy of shared accommodation platform is subsidy. When  $0 < y < y^* < 1$ , the stabilization strategy of shared accommodation platform is nonsubsidy.*

*Proof:* Assume  $G(y) = y * (R_B - \alpha S_B) + z * (R_L - \beta S_L) + H$ ; when  $R_B - \alpha S_B < 0$ ,  $G(y)$  is considered to be a decreasing function of  $y$ . When  $y > y^*$ ,  $G(y) < 0$ ,  $F(x)|_{x=0} = 0$ , and  $F'_x(x)|_{x=0} < 0$ ,  $x=0$  is the stabilization point; that is, the nonsubsidy of the shared accommodation platform is the stabilization strategy; when  $y < y^*$ ,  $G(y) > 0$ ,  $F(x)|_{x=1} = 0$ , and  $F'_x(x)|_{x=1} < 0$ ,  $x=1$  is the stabilization point; that is, the subsidy of the shared accommodation platform is a stabilization strategy; when  $y = y^*$ ,  $x$  is stable in the interval from 0 to 1; that is, regardless of the initial proportion of the strategy selected by the shared accommodation platform, the platform will not change the strategy.

Proposition 1 shows that when the reputational benefits of platform subsidy to consumers are less than the subsidy amount of the platform to consumers, the reduction of consumer participation will change the stability strategy of the shared accommodation platform from nonsubsidy to subsidy. Similarly, the increase of consumer participation will change the stability strategy of the shared accommodation platform from subsidy to nonsubsidy. Therefore, the participation strategy of consumers is crucial to the implementation of platform subsidy, and it is also the basis of platform subsidy strategy.

The phase diagram of the strategic evolution of the shared accommodation platform is shown in Figure 2.

Figure 2 shows that the volume of  $V_1$  represents the proportion of subsidy of the shared accommodation platform, and the volume of  $V_2$  represents the proportion of nonsubsidy of the shared accommodation platform, which is calculated as follows:

$$V_1 = \int_0^1 \int_{-H/(R_L - \beta S_L)}^1 \frac{z * (R_L - \beta S_L) + H}{R_B - \alpha S_B} dz dx = \frac{(H + a_1)^2}{2a_1 * b_1},$$

$$V_2 = 1 - V_1,$$

TABLE 2: Game model benefit matrix of shared accommodation platform, consumer, and landlord.

Landlord	Strategy choice		Shared accommodation platform		
	Consumer	Subsidy ( $x$ )	Subsidy ( $x$ )	Nonsubsidy ( $1-x$ )	
Share ( $z$ )	Participation ( $y$ )	$U_B - C_P + P_B(1-\phi) + P_L + R_B + R_L - \alpha\delta_B - \beta S_L - P_B - C_B + U_B + \alpha\delta_B + N_B\phi P_B - C_P + P_B(1-\phi) + P_L - H - P_B - C_B + U_B + N_B\phi P_B + N_L + U_L - C_L - P_L$ $+ \beta S_L + N_L + U_L - C_L - P_L$	$U_B - C_P + P_B(1-\phi) + P_L + R_B + R_L - \alpha\delta_B - \beta S_L - P_B - C_B + U_B + \alpha\delta_B + N_B\phi P_B - C_P + P_B(1-\phi) + P_L - H - P_B - C_B + U_B + N_B\phi P_B + N_L + U_L - C_L - P_L$ $+ \beta S_L + N_L + U_L - C_L - P_L$	$U_B - C_P + P_B(1-\phi) + P_L + R_B + R_L - \alpha\delta_B - \beta S_L - P_B - C_B + U_B + \alpha\delta_B + N_B\phi P_B - C_P + P_B(1-\phi) + P_L - H - P_B - C_B + U_B + N_B\phi P_B + N_L + U_L - C_L - P_L$ $+ \beta S_L + N_L + U_L - C_L - P_L$	$U_B - C_P + P_B(1-\phi) + P_L + R_B + R_L - \alpha\delta_B - \beta S_L - P_B - C_B + U_B + \alpha\delta_B + N_B\phi P_B - C_P + P_B(1-\phi) + P_L - H - P_B - C_B + U_B + N_B\phi P_B + N_L + U_L - C_L - P_L$ $+ \beta S_L + N_L + U_L - C_L - P_L$
	Nonparticipation ( $1-y$ )	$-C_P + P_L + R_L - \beta S_L - \beta S_L + U_L - C_L - P_L$	$-C_P + P_L + R_L - \beta S_L - \beta S_L + U_L - C_L - P_L$	$-C_P + P_L + R_L - \beta S_L - \beta S_L + U_L - C_L - P_L$	$-C_P + P_L + R_L - \beta S_L - \beta S_L + U_L - C_L - P_L$
Unshare ( $1-z$ )	Participation ( $y$ )	$-C_P + R_B - \alpha\delta_B\alpha\delta_B - C_B A$	$-C_P + R_B - \alpha\delta_B\alpha\delta_B - C_B A$	$-C_P - H - C_B A$	
	Nonparticipation ( $1-y$ )	$-C_P - B A$	$-C_P - B A$	$-C_P - H - B A$	



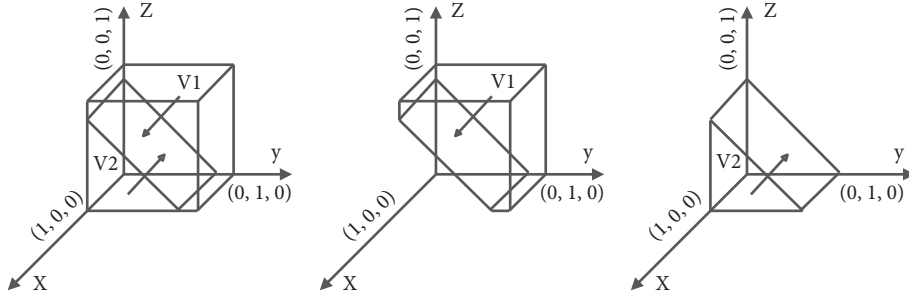


FIGURE 2: Phase diagram of strategy evolution of shared accommodation platforms. (a)  $y = -((z * (R_L - \beta S_L) + H)/(R_B - \alpha S_B))$ , (b)  $y > -((z * (R_L - \beta S_L) + H)/(R_B - \alpha S_B))$ , and (c)  $y < -((z * (R_L - \beta S_L) + H)/(R_B - \alpha S_B))$ .

where  $a_1 = R_L - \beta S_L$  and  $b_1 = R_B - \alpha_1 S_B$ .  $\square$

**Corollary 1.** *The proportion of shared accommodation platform subsidy is positively correlated with the reputation benefits of the platform subsidizing and the opportunity loss of the platform nonsubsidizing and negatively correlated with the subsidy amount of the platform subsidizing to consumers and landlords.*

*Proof:* According to the formula (4), when  $H > \beta S_L - R_L$ ;  $H > \alpha S_B - R_B$ , the first-order partial derivative of each element is obtained as follows:  $(\partial V_1 / \partial R_L) > 0$ ,  $(\partial V_1 / \partial R_B) > 0$ ,  $(\partial V_1 / \partial H) > 0$ ,  $(\partial V_1 / \partial \beta S_L) < 0$ , and  $(\partial V_1 / \partial \alpha S_B) < 0$  (see Appendix A). Therefore, the increase of  $R_L$ ,  $R_B$ , and  $H$  and the decrease of  $\beta S_L$  and  $\alpha S_B$  can increase the proportion of subsidizing for shared accommodation platforms.

Corollary 1 shows that, with the increase in the reputation benefits of the platform due to subsidies to consumers and landlords and the opportunity cost of the platform nonsubsidy, the sharing accommodation platform is more inclined to adopt the subsidy strategy. As the amount of subsidies provided by the platform to landlords and consumers increases, the shared accommodation platform is more inclined to adopt a nonsubsidy strategy.  $\square$

**3.4.2. Stability Analysis of Consumer Strategies.** The expected utility of the consumer choosing to participate is  $\pi_y$ , the expected utility of the consumer choosing nonparticipating is  $\pi_{1-y}$ , then the average expected utility of the consumer is  $\bar{\pi}_y$ , and  $\bar{\pi}_y = y\pi_y + (1-y)\pi_{1-y}$ , which are defined as follows:

$$\begin{aligned} \pi_y &= xz * (-P_B - C_B + U_B + \alpha S_B + N_B) + (1-x)(1-z) * (-C_B) + x(1-z) \\ &\quad * (\alpha S_B - C_B) + (1-x)z * (-P_B - C_B + U_B + N_B), \\ \pi_{1-y} &= xz * (-B) + (1-x)(1-z) * (-B) + x(1-z) * (-B) + (1-x)z * (-B). \end{aligned} \quad (5)$$

According to the Malthusian dynamic equation, the replication dynamic equation of consumer is obtained:

$$\begin{aligned} F(y) &= \frac{dy}{dt} = y(1-y) * [z * (-P_B + N_B) + x * \alpha S_B + U_B - C_B + B], \\ F'_y(y) &= \frac{\partial F(y)}{\partial y} = (1-2y) * [z * (-P_B + N_B) + x * \alpha S_B + U_B - C_B + B]. \end{aligned} \quad (6)$$

According to the stability theorem of the replication dynamic equation, the evolutionary stability strategy of the consumer must be satisfied with the conditions that  $F(y) = 0$  and  $F'_y(y) < 0$ . And the threshold of the consumer strategy transformation is  $z^* = (-x * \alpha S_B + C_B - U_B - B) / (N_B - P_B)$ .

**Proposition 2.** *In the evolution process, when  $0 < z^* < z < 1$ , the stabilization strategy of consumer is participation. When*

$0 < z^* < z < 1$ , the stabilization strategy of consumer is nonparticipation.

*Proof:* Assume  $G(z) = z * (-P_B + N_B) + x * \alpha S_B + U_B - C_B + B$ ; when  $N_B > P_B$ ,  $C_B - U_B - B > \alpha S_B$ ,  $G(z)$  is considered to be an increasing function of  $z$ . When  $z > z^*$ ,  $G(z) > 0$ ,  $F(y)|_{y=1} = 0$ , and  $F'_y(y)|_{y=1} < 0$ , then  $y = 1$  is the stabilization point; that is, consumer participation is the

stabilization strategy; when  $z < z^*$ ,  $G(z) < 0$ ,  $F(y)|_{y=0} = 0$ , and  $F'_y(y)|_{y=0} < 0$ ,  $y = 0$  is the stabilization point; that is, the nonparticipation of the consumer is the stabilization strategy; when  $z = z^*$ ,  $y$  is stable in the interval from 0 to 1; that is, regardless of the initial proportion of the strategy selected by the consumer, the consumer will not change the strategy.

Proposition 2 shows that the strategic choice of the landlord will affect the consumer's choice. When the cross-network externalization utility gained by the consumer with the landlord on the shared platform is more than the fees paid by the consumer participating in the shared accommodation platform, with the increase of the landlord's share, the consumer's stabilization strategy will be changed from nonparticipation to participation. Similarly, the decrease in landlord's share will change the consumer's stabilization strategy from participation to nonparticipation. Therefore, to promote consumer participation in shared accommodation, the platform should take proactive measures to allow more landlords to share their rooms and improve the network cross-externalities utility between consumers and landlords to make it habitual for consumers to participate in shared accommodation.

The phase diagram of the strategic evolution of the consumer is shown in Figure 3.

Figure 3 shows that the volume of  $V_3$  represents the proportion of consumers participating in shared accommodation, and the volume of  $V_4$  represents the proportion of consumers nonparticipation in shared accommodation, which is calculated as follows:

$$V_3 = \int_0^1 \int_{-C_B+U_B+B/\alpha S_B}^1 \frac{-x * \alpha S_B + C_B - U_B - B}{N_B - P_B} dx dy \quad (7)$$

$$= \frac{2a_2 * b_2 - a_2^2 - b_2^2}{2a_2 * c_2},$$

$$V_4 = 1 - V_3, \quad (8)$$

where  $a_2 = \alpha S_B$ ,  $b_2 = C_B - U_B - B$ , and  $c_2 = N_B - P_B$ .

**Corollary 2.** *The proportion of consumer participation in shared accommodation is positively correlated with the subsidy amount given to consumers by the platform, the utility of platform service for the consumer, the fees paid by consumers for renting rooms through offline channels, and the cross-network externalization utility gained by consumers with landlords on the shared platform. The proportion of consumer participation in shared accommodation is negatively correlated with the consumers' time cost and the fees paid by consumers participating in shared accommodation platforms.*

*Proof:* According to the formula (8), when  $C_B > U_B + B$ ;  $C_B - U_B - B > N_B - P_B$ , the first-order partial derivative of each element is obtained as follows:  $(\partial V_3 / \partial \alpha S_B) > 0$ ,  $(\partial V_3 / \partial C_B) < 0$ ,  $(\partial V_3 / \partial U_B) > 0$ ,  $(\partial V_3 / \partial B) > 0$ ,  $(\partial V_3 / \partial N_B) > 0$ , and  $(\partial V_3 / \partial P_B) < 0$  (see Appendix B). Therefore, the increase of  $\alpha S_B$ ,  $U_B$ ,  $B$ , and  $N_B$  and the decrease of  $C_B$  and  $P_B$  can increase the proportion of consumer participation in shared accommodation.

Corollary 2 shows that, with the increase of the subsidies amount for consumers gained from the platform, the utility of platform services for the consumer, the fees paid for obtaining rooms through offline channels, and the utility of cross-network externalities of landlords to consumers, the consumer is more inclined to adopt the participation strategy. As the time cost of consumers and the payment price of consumers' participation in shared accommodation increase, the consumer is more inclined to adopt the nonparticipation strategy.  $\square$

**3.4.3. Stability Analysis of Landlord Strategies.** The expected utility of the landlord choosing to share the rooms is  $\pi_z$ , the expected utility of the landlord choosing to unshare the rooms is  $\pi_{1-z}$ , then the average expected utility of landlord is  $\bar{\pi}_z$ , and  $\bar{\pi}_z = z\pi_z + (1-z)\pi_{1-z}$ , which are defined as follows:

$$\begin{aligned} \pi_z &= xy * (\phi P_B + \beta S_L + N_L + U_L - C_L - P_L) + (1-x)(1-y) * (U_L - C_L - P_L) \\ &\quad + x(1-y) * (\beta S_L + U_L - C_L - P_L) + (1-x)y * [\phi P_B + N_L + U_L - C_L - P_L], \quad (9) \\ \pi_{1-z} &= xy * A + (1-x)y * A + x(1-y) * A + (1-x)(1-y) * A. \end{aligned}$$

According to the Malthusian dynamic equation, the replication dynamic equation of landlord is obtained:

$$F(z) = \frac{dz}{dt} = z(1-z) * [y * (\phi P_B + N_L) + x * \beta S_L + U_L - C_L - P_L - A], \quad (10)$$

$$F'_z(z) = \frac{\partial F(z)}{\partial z} = (1-2z) * [y * (\phi P_B + N_L) + x * \beta S_L + U_L - C_L - P_L - A].$$

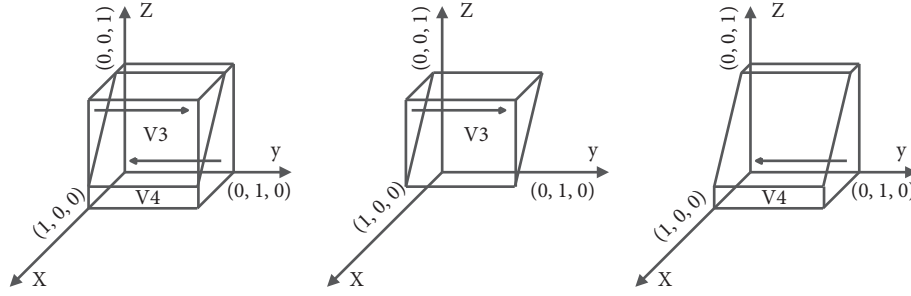


FIGURE 3: Phase diagram of the evolution of the consumer's strategy. (a)  $z = ((-x * \alpha S_B + C_B - U_B - B)/(N_B - P_B))$ , (b)  $z > ((-x * \alpha S_B + C_B - U_B - B)/(N_B - P_B))$ , and (c)  $z < ((-x * \alpha S_B + C_B - U_B - B)/(N_B - P_B))$ .

According to the stability theorem of the replication dynamic equation, the evolutionary stability strategy of the landlord must be satisfied with the conditions that  $F(z) = 0$  and  $F'_z(z) < 0$ . And the threshold of the consumer strategy transformation is  $x^* = (-y * (\phi P_B + N_L) - U_L + C_L + P_L + A)/\beta S_L$ .

**Proposition 3.** *In the evolution process, when  $0 < x^* < x < 1$ , the stabilization strategy of landlord is sharing. When  $0 < x < x^* < 1$ , the stabilization strategy of landlord is unsharing.*

*Proof:* Assume  $G(x) = y * (\phi P_B + N_L) + x * \beta S_L + U_L - C_L - P_L - A$ ; when  $C_L + P_L + A - U_L > \phi P_B + N_L$ ,  $G(x)$  is considered to be an increasing function of  $x$ . When  $x > x^*$ ,  $G(x) > 0$ ,  $F(z)|_{z=1} = 0$ , and  $F'_z(z)|_{z=1} < 0$ ,  $z = 1$  is the stabilization point; that is, the share of the landlord is the stabilization strategy; when  $x < x^*$ ,  $G(x) < 0$ ,  $F(z)|_{z=0} = 0$ , and  $F'_z(z)|_{z=0} < 0$ ,  $z = 0$  is the stabilization point; that is, the

unshare of the landlord is the stabilization strategy; when  $x = x^*$ ,  $z$  is in an evolutionary stabilization state; that is, regardless of the initial proportion of the strategy selected by the landlord, the landlord will not change the strategy.

Proposition 3 shows that the strategy choice of shared accommodation platforms will affect the choice of the landlord. The increase in the proportion of the platform subsidy will change the landlord's stabilization strategy from nonsharing to sharing. Similarly, the reduction of the platform subsidy proportion will change the landlord's stabilization strategy from sharing to nonsharing. Therefore, platform subsidy is a powerful means to promote landlords to share.

The phase diagram of the strategic evolution of the landlord is shown in Figure 4.

Figure 4 shows that the volume of  $V_5$  represents the proportion of the landlord sharing the rooms, and the volume of  $V_6$  represents the proportion of the landlord unsharing the rooms, which is calculated as follows:

$$V_5 = \int_0^1 \int_{-U_L+C_L+P_L+A/\phi P_B+N_L}^1 \frac{-y * (\phi P_B + N_L) - U_L + C_L + P_L + A}{\beta S_L} dy dz = \frac{2a_3 * b_3 - a_3^2 - b_3^2}{2a_3 * c_3}, \quad (11)$$

$$V_6 = 1 - V_5, \quad (12)$$

where  $a_3 = \beta S_L$ ,  $b_3 = -U_L + C_L + P_L + A$ , and  $c_3 = \phi P_B + N_L$ .

**Corollary 3.** *The proportion of the landlord sharing rooms in the shared accommodation platform is positively correlated with the subsidy amount of the platform to landlords, the utility of platform service for the landlord, the fees paid by consumers participating in shared accommodation platforms, and the cross-network externalities utility gained by landlords with consumers on the shared platform. The proportion of the landlord sharing rooms is negatively correlated with the income by the landlord with unsharing room in the platform, the cost of room and service provided by the landlord, and the fees charged to landlords for the platform services.*

*Proof:* According to the formula (11), when  $C_L + P_L + A > U_L$  and  $C_L + P_L + A - U_L > \beta S_L$ , the first-order partial

derivative of each element is obtained:  $(\partial V_5 / \partial \beta S_L) > 0$ ,  $(\partial V_5 / \partial C_L) < 0$ ,  $(\partial V_5 / \partial P_L) < 0$ ,  $(\partial V_5 / \partial A) < 0$ ,  $(\partial V_5 / \partial U_L) > 0$ ,  $(\partial V_5 / \partial \phi P_B) > 0$ , and  $(\partial V_5 / \partial N_L) > 0$  (see Appendix C). Therefore, the increase of  $\beta S_L$ ,  $U_L$ ,  $\phi P_B$ , and  $N_L$  and the decrease of  $A$ ,  $C_L$ , and  $P_L$  can increase the proportion of the landlord participating in the shared rooms.

Corollary 3 shows that, with the increase of subsidy amount of the platform to landlords, the utility of platform service for the landlord, the fees paid by consumers participating in shared accommodation platform, and the utility of cross-network externalities of consumers to landlords, the landlord is more inclined to adopt the share strategy. As the increase of the income by the landlord with unsharing room, the cost of room and service provided by the landlord, and the fees charged to landlords for the platform services, the landlord is more inclined to adopt the nonshare strategy.  $\square$

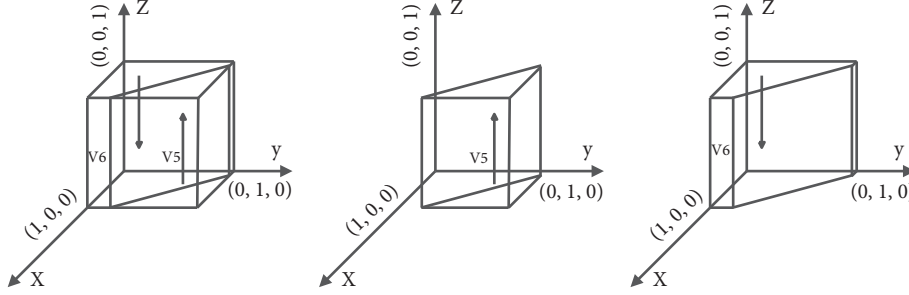


FIGURE 4: Phase diagram of evolution strategy of the landlord. (a)  $x = ((C_L + P_L + A - y * (\phi P_B + N_L) - U_L) / \beta S_L)$ , (b)  $x > ((C_L + P_L + A - y * (\phi P_B + N_L) - U_L) / \beta S_L)$ , and (c)  $x < ((C_L + P_L + A - y * (\phi P_B + N_L) - U_L) / \beta S_L)$ .

## 4. Results and Discussion

**4.1. Strategy Portfolio Stability Analysis.** In the dynamic system of the shared accommodation platform, consumers, and landlords, the stability of the strategic combination of the three-party evolutionary game can be referred to Liapunov's nonlinear function stability discriminant method. When all eigenvalues of the Jacobi matrix are less than zero, the equilibrium point must be asymptotically stabilization; when the first-order bias of the matrix is positive or semipositive and one or two of the eigenvalues of the Jacobi matrix are greater than zero, the equilibrium point is the unstable point or the saddle point. Therefore, in the three-party evolutionary game, the stability of the eight pure strategy equilibrium will be analyzed in this paper. That is,  $E1(0, 0, 0)$ ,  $E2(1, 0, 0)$ ,  $E3(0, 1, 0)$ ,  $E4(0, 0, 1)$ ,  $E5(1, 1, 0)$ ,  $E6(1, 0, 1)$ ,  $E7(0, 1, 1)$ , and  $E8(1, 1, 1)$ .

According to the dynamic equations of each game subject, the Jacobian matrix is obtained as follows:

$$J = \begin{bmatrix} F'_x(x) & F'_y(x) & F'_z(x) \\ F'_x(y) & F'_y(y) & F'_z(y) \\ F'_x(z) & F'_y(z) & F'_z(z) \end{bmatrix}, \quad (13)$$

where the values of the Jacobi matrix are shown in Appendix D and the Jacobian matrix of the equilibrium point is shown in Appendix E.

According to the Jacobian matrix, the stability analysis of the equilibrium point of the three-party evolutionary game can be obtained, and it is shown in Table 3.

When condition I is satisfied, that is  $\alpha S_B + U_B - C_B < B$ ;  $\beta S_L + U_L - C_L - P_L < A$ , the dynamic system has a stable point, that is,  $E2(1, 0, 0)$ . The consumer chooses not to participate in shared accommodation when the sum of the subsidy amount and the net utility gained by consumers from the platform is less than the fees paid by consumers for renting rooms through offline channels. The landlord chooses to unshare the rooms when the sum of the subsidy amount and the net utility gained by the landlord from the platform is less than the income by the landlord with unsharing room in the platform. The stable strategy of sharing accommodation platforms is the subsidy. This may happen in the early stage of the development of shared accommodation. When shared accommodation platform is just getting started, landlords and consumers do not know

enough about shared accommodation. The platform adopts a subsidy strategy to attract landlords and consumers. When the shared accommodation platform has just started, the platform adopts the subsidy strategy to attract landlords and consumers to join in because of their insufficient awareness of shared accommodation.

When condition II is satisfied, that is  $\alpha S_B + \beta S_L - R_B - R_L < H$ ;  $P_B + C_B - B < N_B + \alpha S_B + U_B$ ;  $C_L + P_L + A < \phi P_B + \beta S_L + N_L + U_L$ , the dynamic system has a stable point, that is,  $E8(1, 1, 1)$ . The shared accommodation platform chooses the subsidy strategy when the difference between the total subsidy amount and the reputational benefits of the platform is less than the opportunity loss of platform nonsubsidizing. The consumer chooses to participate in shared accommodation when the difference between the time cost and the fees of consumers in the shared accommodation platform and the fees paid by the consumer for renting rooms through offline channels is less than the utility and benefit that the consumer obtains on the sharing platform. The landlord chooses to share the rooms when the utility and benefit that the landlord obtains on the sharing platform are more than the sum of the cost paid by the landlord on the sharing platform and the income by the landlord with unsharing room in the platform. This may happen during the rapid development of shared accommodation platform. The platform develops appropriate subsidy strategies, such as issuing coupons and cash rewards, to promote more consumers to participate in shared accommodation and landlords to share the rooms, which make landlords and consumers form the habits of participating in sharing accommodation.

When condition III is satisfied, that is  $\alpha S_B + \beta S_L - R_B - R_L > H$ ;  $P_B + C_B - B < N_B + U_B$ ;  $C_L + P_L + A < \phi P_B + N_L + U_L$ , the replicated dynamic system has a stable point, that is,  $E7(0, 1, 1)$ . The shared accommodation platform chooses the subsidy strategy when the difference between the total subsidy amount and the reputational benefits of the platform is more than the opportunity loss of platform nonsubsidizing. The consumer chooses to participate in shared accommodation when the difference between the time cost and the fees of consumers in the shared accommodation platform and the fees paid by the consumer for renting rooms through offline channels is less the utility gained by consumers on the shared platform. The landlord chooses to share the rooms when the utility and benefit that

TABLE 3: Stability analysis of equilibrium point of evolutionary game.

Equilibrium point	Eigenvalue symbol	Stability of the equilibrium point	Condition
$E1(0, 0, 0)$	$(+, +, -)$	Saddle point	\
$E2(1, 0, 0)$	$(-, X, X)$	ESS	I
$E3(0, 1, 0)$	$(+, +, -)$	Saddle point	\
$E4(0, 0, 1)$	$(+, -, +)$	Saddle point	\
$E5(1, 1, 0)$	$(+, +, -)$	Unstable point	\
$E6(1, 0, 1)$	$(-, X, +)$	Unstable point	\
$E7(0, 1, 1)$	$(X, X, X)$	ESS	III
$E8(1, 1, 1)$	$(X, X, X)$	ESS	II

Note.  $X$  means uncertainty of symbol, and ESS means the evolutionarily stable strategy. Condition I:  $\alpha S_B + U_B - C_B < B$ ;  $\beta S_L + U_L - C_L - P_L < A$ . Condition II:  $\alpha S_B + \beta S_L - R_B - R_L < H$ ;  $P_B + C_B - B < N_B + \alpha S_B + U_B$ ;  $C_L + P_L + A < \phi P_B + \beta S_L + N_L + U_L$ . Condition III:  $\alpha S_B + \beta S_L - R_B - R_L > H$ ;  $P_B + C_B - B < N_B + U_B$ ;  $C_L + P_L + A < \phi P_B + N_L + U_L$ .

the landlord obtains on the sharing platform are more than the sum of the cost paid by the landlord on the sharing platform and the income by the landlord with unsharing room in the platform. This may happen in the mature stage of shared accommodation platform development. The landlords and the consumers have already understood shared accommodation, and the platform has developed a series of sound policies to push the information and serve the landlords and consumers. The platforms can make considerable gains by choosing nonsubsidizing.

**4.2. Simulation Analysis.** In order to test the reliability of the model and to more intuitively observe the evolution trajectory of stakeholders and their sensitivity to parameters, this paper uses MATLAB2016 to simulate the evolution stability strategy and related parameters.

**4.2.1. Model Testing.** When condition I is satisfied, the simulation parameters are set as follows:  $C_P = 5$ ,  $C_B = 11$ ,  $C_L = 3$ ,  $R_B = R_L = 2$ ,  $S_B = S_L = 5$ ,  $P_L = P_B = 5$ ,  $\alpha = 0.1$ ,  $\beta = 0.3$ ,  $\phi = 0.7$ ,  $A = H = 3$ ,  $B = 6$ ,  $N_B = 6$ ,  $N_L = 2$ ,  $U_B = 1$ , and  $U_L = 3$ . Selecting four different initial strategies, all four groups of initial strategies all evolve towards  $E2(1, 0, 0)$ , and the results of simulation and evolution with time are shown in Figure 5.

It can be seen from Figure 5 that when shared accommodation platform is in the early stage of development, the consumer and landlord do not have a certain understanding of the sharing accommodation platform. For the needs of market expansion, shared accommodation platform subsidizes consumers and landlords to promote them to join the shared accommodation platform.

When condition II is satisfied, the simulation parameters are set as follows:  $C_B = 8$ ,  $P_L = 3$ ,  $\alpha = 0.6$ , and  $\beta = 0.5$ , and other parameters are the same with the equilibrium point  $E2(1, 0, 0)$ . Selecting four different initial strategies, all four groups of initial strategies all evolve towards  $E8(1, 1, 1)$ , and the results of simulation and evolution with time are shown in Figure 6.

It can be seen from Figure 6 that, with the reduction of the time cost of the consumer participating in shared accommodation and the increase of the service utility and the subsidies to consumers by the platform, the actual cost of consumer participating in shared accommodation is less

than the cost through offline channels. Therefore, the consumer will choose to participate in shared accommodation platform. With the increase of the subsidy to the landlord and reduction of the service fee for the landlord to participate in the shared accommodation platform, the landlord's revenue from shared rooms is more than the loss of the landlord's inventory for nonparticipating in sharing rooms. The landlord will adopt the strategy to actively join the shared accommodation platform to share rooms. In the end, the platform, consumers, and landlords will tend to adopt the positive strategies.

When condition III is satisfied, the simulation parameters are set as follows:  $C_L = 1$ ,  $P_B = 3$ ,  $\phi = 0.9$ , and  $H = 1$ , and other parameters are the same with the equilibrium point  $E8(1, 1, 1)$ . Selecting four different initial strategies, all four groups of initial strategies all evolve towards  $E7(0, 1, 1)$ , and the results of simulation and evolution with time are shown in Figure 7.

It can be seen from Figure 7 that, with the reduction of the time cost for consumers and the fees of the landlords to participate in shared accommodation, the consumer and landlord are promoted to participate in the shared accommodation platform. When the consumers and landlords become accustomed to participating in the shared accommodation platform to reduce the opportunity loss of platform nonsubsidizing, the platform can appropriately reduce the subsidies and the cost paid by consumers to participate in shared accommodation and increase the share of the revenue for landlords. Consumers and landlords will still choose to participate in shared accommodation. In the mature stage of the development of the shared accommodation platform, the platform can set the reasonable price for consumers and the preferential proportion of the landlord's share in expenses under the nonsubsidy policy, the landlords, and consumers can continue to participate in the shared accommodation to maximize the interests of all three parties.

**4.2.2. The Influence of Different Subjects' Initial Willingness to Participate.** The initial proportions of the strategies of the shared accommodation platform, the consumer, and the landlord are set as follows:  $(0.2, 0.2, 0.2)$ ,  $(0.5, 0.5, 0.5)$ , and  $(0.9, 0.9, 0.9)$ . The stability of the system evolution is shown in Figure 8.



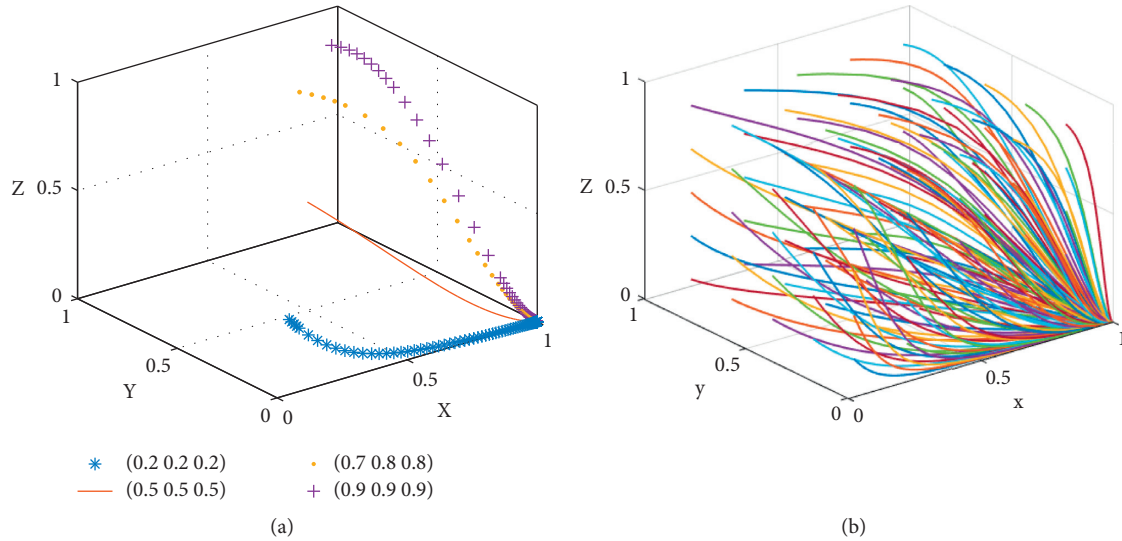


FIGURE 5: Evolutionary equilibrium diagram at the point  $E2(1, 0, 0)$ .

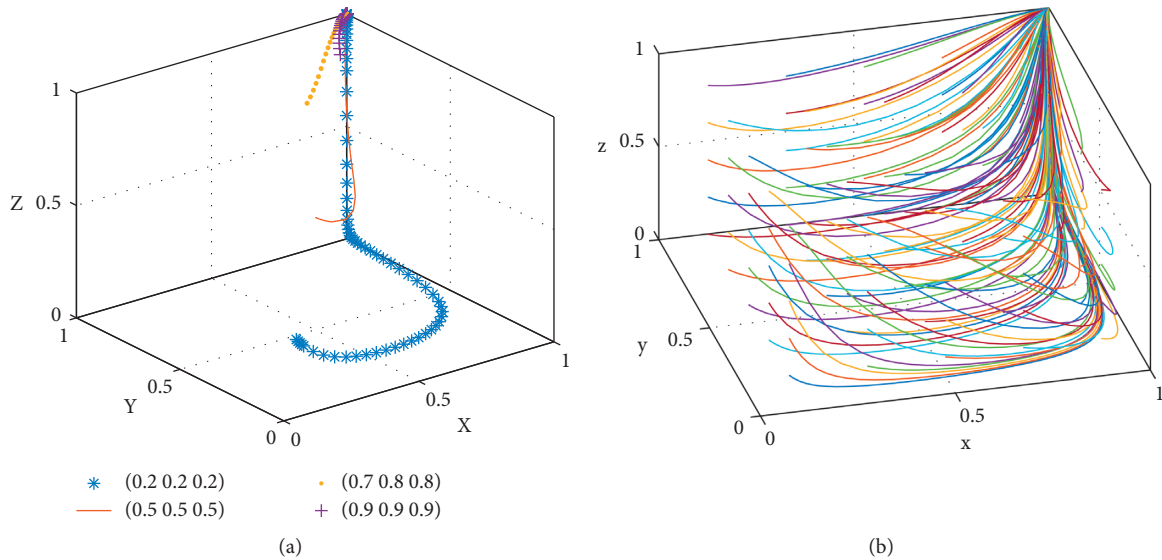


FIGURE 6: Evolutionary equilibrium diagram at the point  $E8(1, 1, 1)$ .

It can be seen from Figure 8 that, under the parameter conditions of the evolutionary stability strategy  $E8(1, 1, 1)$ , the behavioral strategies of the shared accommodation platform, consumers, and landlords finally evolve into the strategies of “subsidy,” “participation,” and “share” with the initial proportions of the strategies of  $(0.2, 0.2, 0.2)$ ,  $(0.5, 0.5, 0.5)$ , and  $(0.9, 0.9, 0.9)$ . This indicates that the stable state of the system is independent of the initial proportions of the strategies. Further analysis reveals that, in the case of low initial willingness to participate, landlords choose the unsharing strategy because they can obtain the rental income of the rooms in other ways. When the platform’s subsidy for consumers and landlords increases and the time cost of consumers and landlords’ service fees to participate in shared accommodation platform decreases, the consumer chooses to participate, and landlord will choose to share. When the initial willingness of the platform, consumer,

and landlord is high, all three participants will continue to evolve steadily to the strategies of “subsidy,” “participation,” and “share.” The closer the initial proportion is to the proportion of an evolutionary stable state, the shorter the time for evolution to reach the stable state is.

*4.2.3. The Influence of the Proportion of Subsidies Provided by the Platform to Consumers.* If  $\alpha = \{0.1, 0.4, 0.6\}$ , the stabilization strategy of each participant is obtained as shown in Figure 9.

From Figure 9, it can be seen that, with the increase of the proportion of subsidy to the consumer, the evolutionary stability strategy transitions from  $E2(1, 0, 0)$  to  $E8(1, 1, 1)$ . The critical value of the proportion of subsidy to consumers is between 0.4 and 0.6. When the proportion of subsidies to the



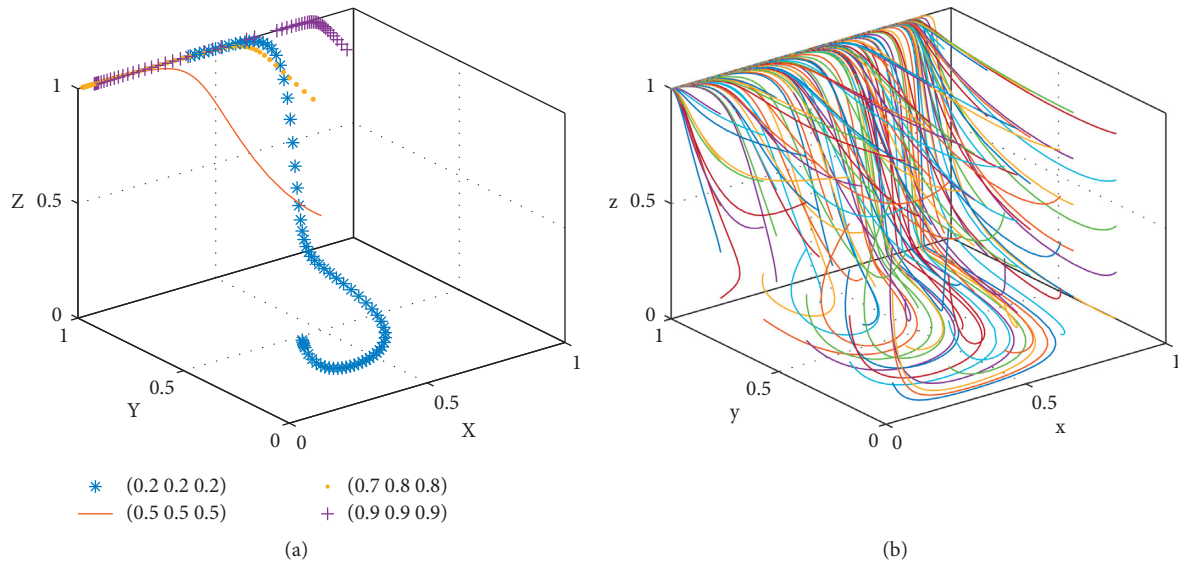


FIGURE 7: Evolutionary equilibrium diagram at the point  $E7(0, 1, 1)$ .

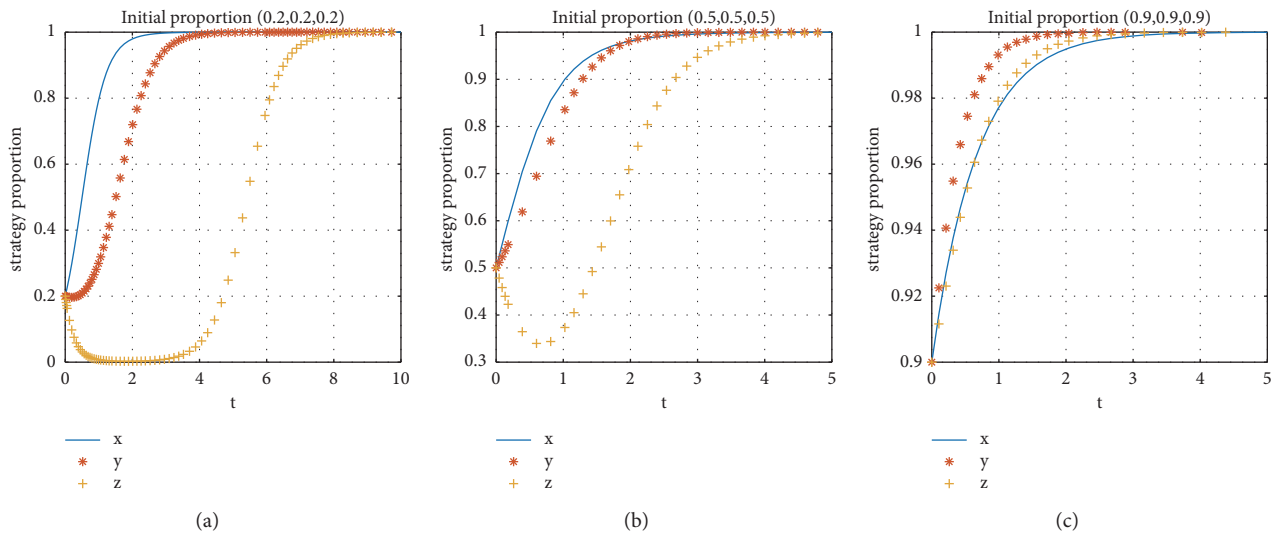


FIGURE 8: Influence of different subjects' initial willingness to participate on the evolutionary.

consumer is larger than the critical value, the evolutionary stability strategy of the three-party game converges to  $E8(1, 1, 1)$ , and the larger the proportion of consumer subsidies is, the faster the evolutionary stability strategy converges. When the proportion of consumer subsidies is less than the critical value, consumer chooses not to participate in the shared accommodation platform, the landlords choose to unshare the rooms, and the time for the landlord to reach stability is faster than that for consumers. Therefore, the platform can increase the proportion of subsidy for consumers to promote the participation of the consumer and the landlord.

**4.2.4. The Influence of the Proportion of Subsidies Provided by the Platform to Landlords.** If  $\beta = \{0.3, 0.5, 0.6\}$ , the stabilization strategy of each participant is obtained as shown in Figure 10.

From Figure 10, it can be seen that, with the increase of the proportion of the subsidy to the landlord, the evolutionary stability strategy transitions from  $E2(1, 0, 0)$  to  $E8(1, 1, 1)$ . The critical value of the proportion of the subsidy to the landlord is between 0.5 and 0.6. When the proportion of subsidies to the landlord is larger than the critical value, the evolutionary stability strategy of the three-party game converges to  $E8(1, 1, 1)$ , and the larger the proportion of landlord subsidies is, the faster the evolutionary stability strategy converges. When the proportion of subsidies to the landlord is less than the critical value, the consumer chooses not to participate in the shared accommodation, the landlord chooses to unshare the rooms, and the time for the landlord to reach stability is slower than that for the consumer to reach stability. It is a benefit for landlords to participate in the shared accommodation platform to increase the subsidies to landlords. Therefore, in the early stage of the platform's development, choosing the more

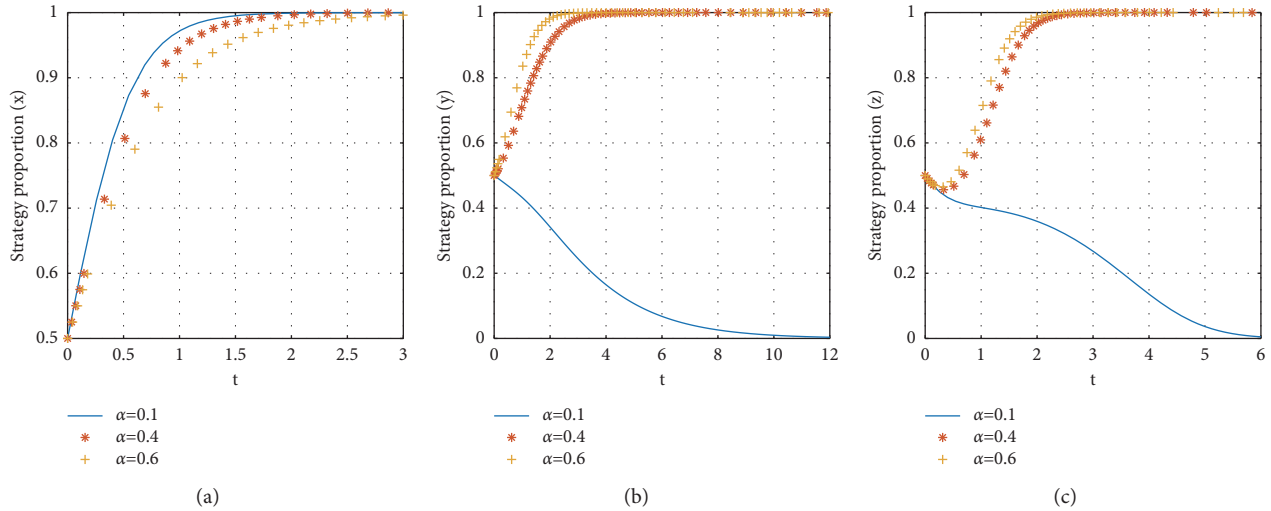


FIGURE 9: The influence of the proportion of subsidies to the consumer on evolutionary stabilization strategies of subject behaviors.

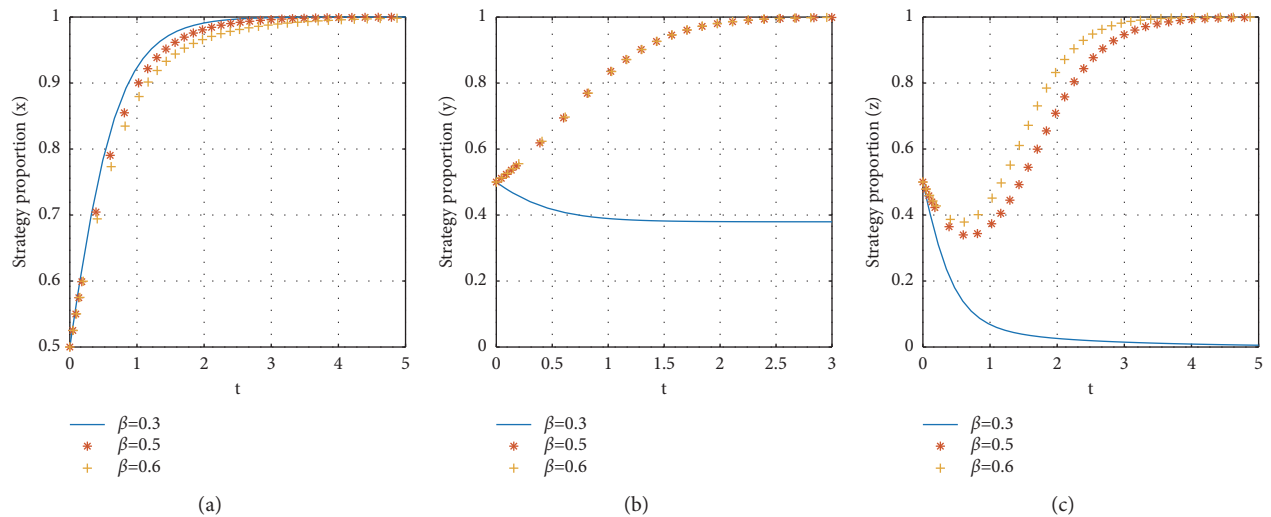


FIGURE 10: The influence of the proportion of subsidies to the landlord on evolutionary stabilization strategies of subject behaviors.

share of subsidy to the landlord can promote consumers and landlords to join the shared accommodation platform more quickly.

4.2.5. *The Influence of the Fees Charged to Landlords for the Platform Services.* If  $P_L = \{3, 4, 5\}$ , the stabilization strategy of each participant is obtained as shown in Figure 11.

From Figure 11, it can be seen that, with the decrease of the fees charged to landlords for the platform services, the evolutionary stability strategy transitions from  $E2(1, 0, 0)$  to  $E8(1, 1, 1)$ . The critical value of the fees charged to landlords for the platform services is between 4 and 5. When the fees to the landlord for the platform services are less than the critical value, the evolutionary stability strategy of the three-party game converges to  $E8(1, 1, 1)$ , and the less the fees to the

landlord for the platform services are, the faster the evolutionary stability strategy converges. When the fees to the landlord for the platform services are more than the critical value, the consumer chooses not to participate in shared accommodation, the landlord chooses to unshare the rooms, and the time for the landlord to reach stability is slower than that for the consumer to reach stability. Therefore, in the early stage of the platform’s development, charging fewer fees to landlords for the platform services can promote consumers and landlords to join the shared accommodation platform more quickly.

4.2.6. *The Influence of the Time Costs of the Consumer.* If  $C_B = \{8, 9, 10\}$ , the stabilization strategy of each participant is obtained as shown in Figure 12.

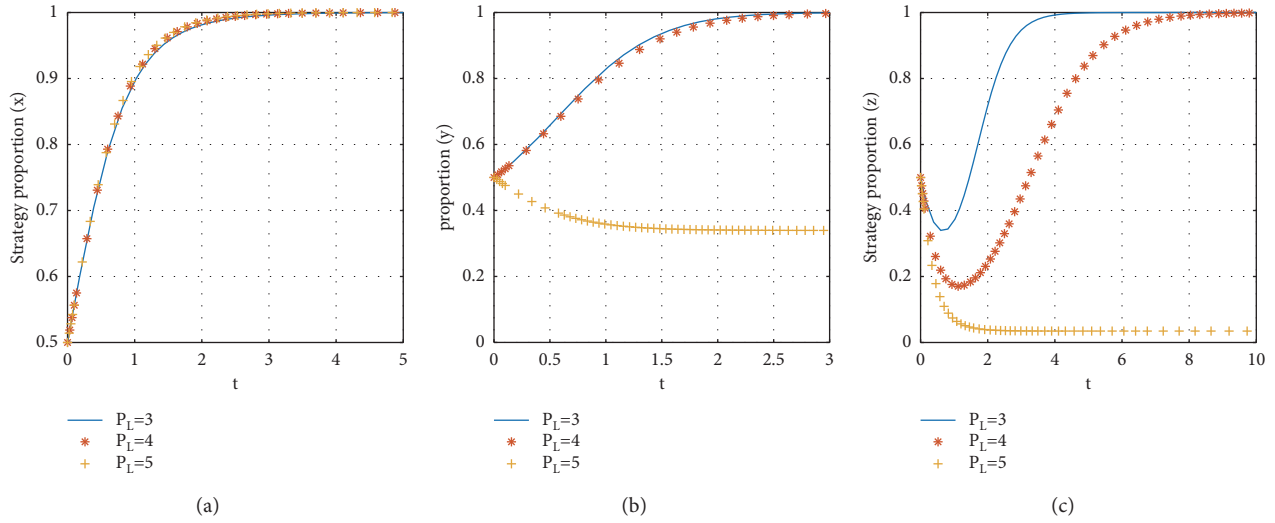


FIGURE 11: The influence of the fees to the landlord for the platform services on evolutionary stabilization strategies of subject behaviors.

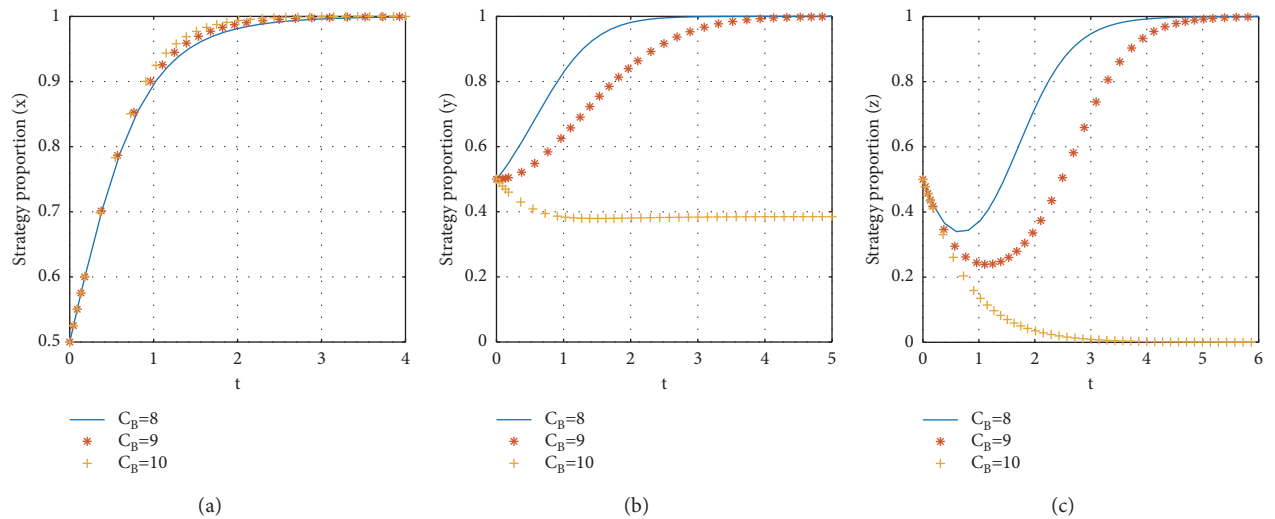


FIGURE 12: The influence of the consumer's time costs on evolutionary stabilization strategies of subject behaviors.

From Figure 12, it can be seen that, with the decrease of the consumer's time costs, the evolutionary stability strategy transitions from  $E2(1, 0, 0)$  to  $E8(1, 1, 1)$ . The critical value of the time cost of the consumer to participate in the shared accommodation platform is between 9 and 10. When the time costs of the consumer are less than the critical value, the evolutionary stability strategy of the three-party game converges to  $E8(1, 1, 1)$ , and the less the consumer's time costs are, the faster the evolutionary stability strategy converges. When the consumer's time costs are more than the critical value, the consumer chooses not to participate in the shared accommodation, the landlord chooses to unshare the rooms, and the time for the landlord to reach stability is slower than the time for the consumer to reach stability. Therefore, in the early stage of the platform's development, decreasing the time costs of the consumer can promote

consumers and landlords to join the shared accommodation platform more quickly.

4.2.7. *The Influence of the Proportion of Landlord's Share in Expenses Paid by Consumers.* If  $\phi = \{0.7, 0.8, 0.9\}$ , the stabilization strategy of each participant is obtained as shown in Figure 13.

From Figure 13, it can be seen that, with the increase of the proportion of landlords' share in expenses, the evolutionary stability strategy transitions from  $E8(1, 1, 1)$  to  $E7(0, 1, 1)$ . The critical value of the proportion of the landlord's share in expenses paid by consumers is between 0.7 and 0.8. When the proportion of the landlord's share in expenses is less than the critical value, the evolutionary stability strategy of the three-party game converges to  $E8(1, 1, 1)$ . When the value of the proportion

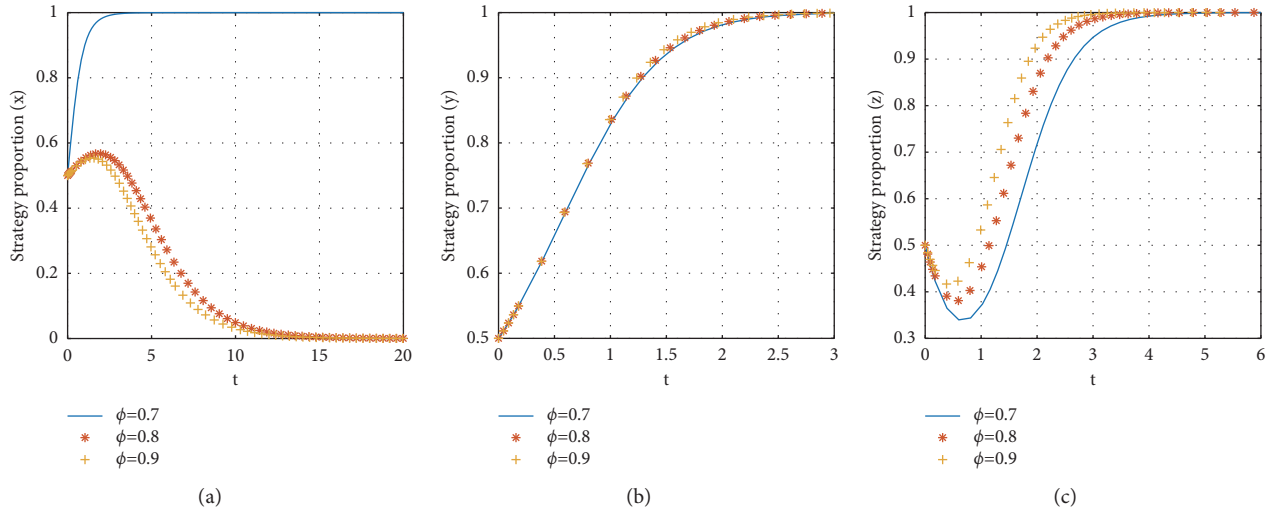


FIGURE 13: The influence of the proportion of landlord’s share in expenses on evolutionary stabilization strategies of subject behaviors.

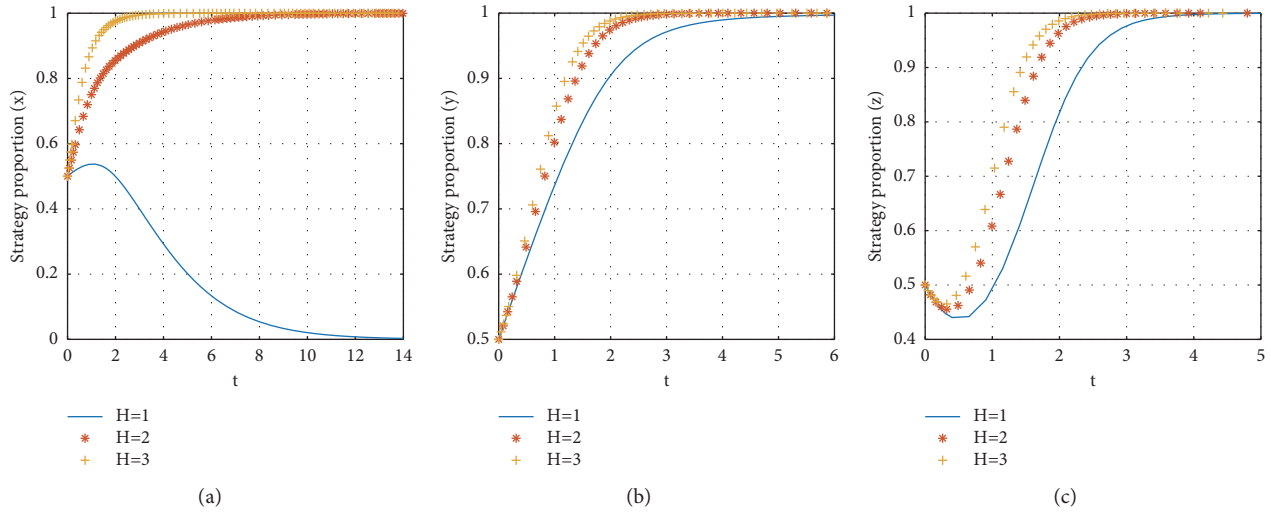


FIGURE 14: The influence of the platform’s opportunity loss on evolutionary stabilization strategies of subject behaviors.

of the landlord’s share in expenses is larger than the critical value, the shared accommodation platform chooses the non-subsidy strategy. And the larger the proportion of the landlord’s share in expenses is, the faster the evolutionary stability strategy of the shared accommodation platform converges. Therefore, in the rapid development of shared accommodation platforms, increasing the proportion of landlords’ share in expenses to meet the interests of landlords can stabilize the choice of the landlord participating in the shared accommodation platform. When the net income of landlords and consumers reaches a certain level with the increase of the scale, the shared accommodation platform can achieve the optimization of system benefits with nonsubsidizing.

4.2.8. *The Influence of the Opportunity Loss of Platform Nonsubsidizing.* If  $H = \{1, 2, 3\}$ , the stabilization strategy of each participant is obtained as shown in Figure 14.

From Figure 14, it can be seen that, with the decrease of the opportunity loss of platform nonsubsidizing, the evolutionary stability strategy transitions from  $E8(1, 1, 1)$  to  $E7(0, 1, 1)$ . The critical value of the opportunity loss of platform nonsubsidizing is between 1 and 2. When the opportunity loss of platform nonsubsidizing is more than the critical value, the evolutionary stability strategy of the three-party game converges to  $E8(1, 1, 1)$ . And when the opportunity loss of platform nonsubsidizing is less than the critical value, the shared accommodation platform chooses the nonsubsidy strategy. The less the opportunity loss of platform nonsubsidizing is, the faster the evolutionary stability strategy of the shared accommodation platform converges. Therefore, in the rapid development of shared accommodation platforms, the decrease of the opportunity loss of platform nonsubsidizing indicates that the consumer and landlord have formed the habit of participating in shared accommodation platform, and the

strategy of the platform gradually evolved to the nonsubsidy.

## 5. Conclusions

Internet and big data technologies provide the technological impetus for the development of new economic formats, and effective governance has become the key to the success of platforms [51, 52]. Platform-based companies aim to maximize the revenue of the platform, and gaining economic benefits is also the important intention of each subject to participate in collaborative consumption [53]. Aiming at the subsidy strategy of shared accommodation platform, an evolutionary game model was established, the system stability was analyzed and simulated, and the main conclusions of this paper are as follows:

- (1) The evolutionary stability strategy of the shared accommodation platform based on subsidizing can go through three stages. At the early stage of platform establishment, the shared accommodation platform attracts consumers and landlords to join through subsidizing due to the need for market expansion. At this time, consumers refuse to join the shared accommodation platform for the high time cost, landlords refuse to join for the high service fee paid, and the system will stabilize on the strategic combination of  $E2$ (subsidy, nonparticipation, unsharing). In this state, consumers and landlords have a preliminary understanding and experience of the sharing accommodation platform. When the difference between the time cost and the fees of consumers in the shared accommodation platform and the fees paid by the consumer for renting rooms through offline channels is less than the utility and benefit that the consumer obtains on the sharing platform, the consumer chooses to join the shared accommodation platform. When the utility and benefit that the landlord obtains on the sharing platform are more than the sum of the cost paid by the landlord on the sharing platform and the income by the landlord with unsharing room in the platform, the landlord chooses to join in the shared accommodation platform. At this time, the system stabilizes on the strategic combination of  $E8$ (subsidy, participation, sharing). Due to the influx of landlords and consumers, the shared accommodation platform is expanding rapidly. When the difference between the total subsidy amount and the reputational benefits of the platform is more than the opportunity loss of platform nonsubsidizing, the platform chooses the nonsubsidy strategy. At this time, it is necessary to meet that the utility of consumers and landlords is more than their costs, and the system will stabilize on the strategic combination of  $E7$ (nonsubsidy, participation, sharing). The shared

accommodation platform no longer needs subsidy to attract landlords and consumers and relies on quality services to achieve good operation of the platform, and the shared accommodation platform reaches the mature stage of development. The dynamic change analysis of the strategies in these three stages makes up for the deficiency of the subsidy strategy of the accommodation platform in the existing research.

- (2) When the strategy combination of the three parties transitions from  $E2$ (subsidy, nonparticipation, unshare) to  $E8$ (subsidy, participation, share), it can be achieved by increasing the subsidies to consumers and landlords and reducing consumers' time costs and service fees charged to the landlord. These measures to increase the profits of consumers and landlords and to reduce their losses are all aimed at attracting consumers and landlords to join the shared accommodation platform. This is consistent with the conclusion proposed by Guo et al. [27] that the bilateral platform of the sharing economy can adopt the subsidy strategy.
- (3) When the strategy combination of the three parties transitions from  $E8$ (subsidize, participate, share) to  $E7$ (nonsubsidy, participation, share), it can be achieved by reducing the opportunity loss of platform nonsubsidy and increasing the proportion of the landlord's share in expenses paid by consumers. Therefore, as consumers and landlords form the habit of participating in shared accommodation platform, landlords and consumers can maintain good profits for the three parties through normal transactions so that the loss of the platform nonsubsidy is reduced, and the strategy of the platform is gradually changed from subsidy to nonsubsidy.

In this paper, the evolutionary game model of three parties in the shared accommodation platform is established, and it is found that the subsidy strategy of the shared accommodation platform can encourage consumers and landlords to participate in shared accommodation. By adjusting the subsidies to customers and landlords, the time cost of consumers, the service fees of landlords, the proportion of landlord's share of consumers' fees, and the loss of the platform's nonsubsidy, the evolutionary equilibrium of the game strategy can be changed. This study makes up for the deficiency of three-party game analysis in the existing literature and does in-depth research on sharing economy by using the game theory.

In order to activate the sharing market and further develop the sharing economy platform, the following suggestions are put forward:

- (1) The subsidies can be used by the platform to attract landlords and consumers. When the platform is in the initial development stage, the subsidies are an effective way to expand the number of platform users. However, the specific subsidy amount must be



considered in combination with the use costs of consumers and landlords, and the platform will try to give an attractive subsidy amount. If the construction of the platform goes smoothly, the subsidy policy can also be adjusted to realize the normal operation of the shared accommodation platform.

- (2) In the development process of the platform, by fully excavating the massive and scattered rooms, the diversity of shared accommodation products is ensured to provide consumers with diverse accommodation options. The platform can use information integration to optimize the platform interface for reducing the time for consumers to choose room information. On this basis, the standardization of shared accommodation services is promoted to further improve the quality of products and services.
- (3) The service utility of the platform to the landlords can be increased to maintain the stability of the landlord's rooms in order to ensure the supply for customers. The platform can reduce the service fee charged to the landlord and increase services quality for landlords. For example, Airbnb provides photo services to landlords, smart door locks, face-brushing check-in, and so on. The platform can also establish a quick communication channel between landlords and customers to help landlords reduce communication costs.

This study establishes a three-party game evolution model based on subsidy policies. Due to the limitations of model assumptions, under the condition of asymmetric information and bounded rationality, the benefits of shared accommodation platforms, consumers, and landlords are maximized in subsidy policies, and the possibility of unsuccessful matching between consumers and landlords in the shared accommodation platform is not considered. Therefore, in order to improve the operating mechanism of the shared accommodation platform, it will be the further research direction to consider the role of factors affecting participants' participation in the strategic evolution, such as the possibility of matching between consumers and landlords, as well as the location and facility differences of rooms.

## Appendix

### A. Supplementary Proof of Corollary 1

The proportion of subsidy strategy of shared accommodation platform is  $V_1$ , and the first-order partial derivative of each element is proved as follows:

$$V_1 = \frac{(H + a_1)^2}{2a_1 * b_1}, \quad (\text{A.1})$$

where  $a_1 = R_L - \beta S_L$  and  $b_1 = R_B - \alpha_1 S_B$ .

When  $b_1 < 0$  and  $H + a_1 > 0$ ,

$$\begin{aligned} \frac{\partial V_1}{\partial R_L} &= \left( \frac{(H + a_1)^2}{2a_1 * b_1} \right)' = -\frac{H(H + a_1)}{a_1^2 * b_1} > 0, \\ \frac{\partial V_1}{\partial R_B} &= \left( \frac{(H + a_1)^2}{2a_1 * b_1} \right)' = -\frac{(H + a_1)^2}{2a_1 * b_1^2} > 0, \\ \frac{\partial V_1}{\partial H} &= \left( \frac{(H + a_1)^2}{2a_1 * b_1} \right)' = \frac{H + a_1}{a_1 * b_1} > 0, \\ \frac{\partial V_1}{\partial \beta S_L} &= \left( \frac{(H + a_1)^2}{2a_1 * b_1} \right)' = -\frac{\partial V_1}{\partial R_L} < 0, \\ \frac{\partial V_1}{\partial \alpha S_B} &= \left( \frac{(H + a_1)^2}{2a_1 * b_1} \right)' = -\frac{\partial V_1}{\partial R_B} < 0. \end{aligned} \quad (\text{A.2})$$

### B. Supplementary Proof of Corollary 2

The proportion of participating strategy of the consumer is  $V_3$ , and the first-order partial derivative of each element is proved as follows:

$$V_3 = \frac{2a_2 * b_2 - a_2^2 - b_2^2}{2a_2 * c_2}, \quad (\text{B.1})$$

where  $a_2 = \alpha S_B$ ,  $b_2 = C_B - U_B - B$ , and  $c_2 = N_B - P_B$ .

$$\begin{aligned} \frac{\partial V_3}{\partial \alpha S_B} &= \left( \frac{2a_2 * b_2 - a_2^2 - b_2^2}{2a_2 * c_2} \right)' = \frac{b_2^2 - 2b_2 * c_2 - a_2^2}{2a_2^2 * c_2} > 0, \\ \frac{\partial V_3}{\partial C_B} &= \left( \frac{2a_2 * b_2 - a_2^2 - b_2^2}{2a_2 * c_2} \right)' = \frac{c_2 - b_2}{a_2 * c_2} < 0, \\ \frac{\partial V_3}{\partial U_B} &= \frac{\partial V_3}{\partial B} = \left( \frac{2a_2 * b_2 - a_2^2 - b_2^2}{2a_2 * c_2} \right)' = -\frac{\partial V_3}{\partial C_B} > 0, \\ \frac{\partial V_3}{\partial N_B} &= \left( \frac{2a_2 * b_2 - a_2^2 - b_2^2}{2a_2 * c_2} \right)' = \frac{a_2^2 + b_2^2}{2a_2 * c_2^2} > 0, \\ \frac{\partial V_3}{\partial P_B} &= \left( \frac{2a_2 * b_2 - a_2^2 - b_2^2}{2a_2 * c_2} \right)' = -\frac{\partial V_3}{\partial N_B} < 0. \end{aligned} \quad (\text{B.2})$$

### C. Supplementary Proof of Corollary 3

The proportion of share strategy of the landlord is  $V_5$ , and the first-order partial derivative of each element is proved as follows:

$$V_5 = \frac{2a_3 * b_3 - a_3^2 - b_3^2}{2a_3 * c_3}, \quad (\text{C.1})$$

where  $a_3 = \beta S_L$ ,  $b_3 = -U_L + C_L + P_L + A$ , and  $c_3 = \phi P_B + N_L$ .



$$\frac{\partial V_5}{\partial \beta S_L} = \left[ \frac{2a_3 * b_3 - a_3^2 - b_3^2}{2a_3 * c_3} \right]' = \frac{b_3^2 - 2b_3 * c_3 - a_3^2}{2a_3^2 * c_3} > 0,$$

$$\frac{\partial V_5}{\partial C_L} = \left[ \frac{2a_3 * b_3 - a_3^2 - b_3^2}{2a_3 * c_3} \right]' = \frac{c_3 - b_3}{a_3 * c_3} < 0,$$

$$\frac{\partial V_5}{\partial P_L} = \frac{\partial V_5}{\partial A} = \left[ \frac{2a_3 * b_3 - a_3^2 - b_3^2}{2a_3 * c_3} \right]' = \frac{\partial V_5}{\partial C_L} < 0,$$

$$\frac{\partial V_5}{\partial U_L} = \left[ \frac{2a_3 * b_3 - a_3^2 - b_3^2}{2a_3 * c_3} \right]' = -\frac{\partial V_5}{\partial C_L} > 0,$$

$$\frac{\partial V_5}{\partial \phi P_B} = \left[ \frac{2a_3 * b_3 - a_3^2 - b_3^2}{2a_3 * c_3} \right]' = \frac{a_3^2 + b_3^2}{2a_3 * c_3} > 0,$$

$$\frac{\partial V_5}{\partial N_L} = \frac{2a_3 * b_3 - a_3^2 - b_3^2}{2a_3 * c_3} = \frac{\partial V_5}{\partial \phi P_B} > 0.$$

(C.2)

## D. Jacobian Matrix of Replicator Dynamic System

The elements of the Jacobian matrix is as follows:

$$F(x) = \frac{dx}{dt} = x(1-x) * [y * (R_B - \alpha S_B) + z * (R_L - \beta S_L) + H],$$

$$F'_x(x) = \frac{\partial F(x)}{\partial x} = (1-2x) * [y * (R_B - \alpha S_B) + z * (R_L - \beta S_L) + H],$$

$$F'_y(x) = \frac{\partial F(x)}{\partial y} = x(1-x) * (R_B - \alpha S_B),$$

$$F'_z(x) = \frac{\partial F(x)}{\partial z} = x(1-x) * (R_L - \beta S_L),$$

$$F(y) = \frac{dy}{dt} = y(1-y) * [z * (-P_B + N_B) + x * \alpha S_B + U_B - C_B + B],$$

$$F'_x(y) = \frac{\partial F(y)}{\partial x} = y(1-y) * \alpha S_B,$$

$$F'_y(y) = \frac{\partial F(y)}{\partial y} = (1-2y) * [z * (-P_B + N_B) + x * \alpha S_B + U_B - C_B + B],$$

$$F'_z(y) = \frac{\partial F(y)}{\partial z} = y(1-y) * (-P_B + N_B),$$

$$F(z) = \frac{dz}{dt} = z(1-z) * [y * (\phi P_B + N_L) + x * \beta S_L + U_L - C_L - P_L - A],$$

$$F'_x(z) = \frac{\partial F(z)}{\partial x} = z(1-z) * \beta S_L,$$

$$F'_y(z) = \frac{\partial F(z)}{\partial y} = z(1-z) * (\phi P_B + N_L),$$

$$F'_z(z) = \frac{\partial F(z)}{\partial z} = (1-2z) * [y * (\phi P_B + N_L) + x * \beta S_L + U_L - C_L - P_L - A].$$

(D.1)

### E. Jacobian Matrix of Equilibrium Points of Replicator Dynamic System

The Jacobian matrix of the equilibrium point  $E1(0, 0, 0)$  is as follows:

$$J_1 = \begin{bmatrix} H & 0 & 0 \\ 0 & U_B - C_B + B & 0 \\ 0 & 0 & U_L - C_L - P_L - A \end{bmatrix}. \quad (E.1)$$

The Jacobian matrix of the equilibrium point  $E2(1, 0, 0)$  is as follows:

$$J_2 = \begin{bmatrix} -H & 0 & 0 \\ 0 & \alpha S_B + U_B - C_B + B & 0 \\ 0 & 0 & \beta S_L + U_L - C_L - P_L - A \end{bmatrix}. \quad (E.2)$$

The Jacobian matrix of the equilibrium point  $E3(0, 1, 0)$  is as follows:

$$J_3 = \begin{bmatrix} R_B - \alpha S_B + H & 0 & 0 \\ 0 & -U_B + C_B - B & 0 \\ 0 & 0 & \phi P_B + N_L + U_L - C_L - P_L - A \end{bmatrix}. \quad (E.3)$$

The Jacobian matrix of the equilibrium point  $E4(0, 0, 1)$  is as follows:

$$J_4 = \begin{bmatrix} R_L - \beta S_L + H & 0 & 0 \\ 0 & U_B - C_B + B - P_B + N_B & 0 \\ 0 & 0 & -U_L + C_L + P_L + A \end{bmatrix}. \quad (E.4)$$

The Jacobian matrix of the equilibrium point  $E5(1, 1, 0)$  is as follows:

$$J_5 = \begin{bmatrix} -R_B + \alpha S_B - H & 0 & 0 \\ 0 & -\alpha S_B - U_B + C_B - B & 0 \\ 0 & 0 & \phi P_B + N_L + \beta S_L + U_L - C_L - P_L - A \end{bmatrix}. \quad (E.5)$$

The Jacobian matrix of the equilibrium point  $E6(1, 0, 1)$  is as follows:

$$J_6 = \begin{bmatrix} -R_L + \beta S_L - H & 0 & 0 \\ 0 & -P_B + N_B + \alpha S_B + U_B - C_B + B & 0 \\ 0 & 0 & -\beta S_L - U_L + C_L + P_L + A \end{bmatrix}. \quad (E.6)$$

The Jacobian matrix of the equilibrium point  $E7(0, 1, 1)$  is as follows:

$$J_7 = \begin{bmatrix} R_B - \alpha S_B + R_L - \beta S_L + H & 0 & 0 \\ 0 & P_B - N_B - U_B + C_B - B & 0 \\ 0 & 0 & -\phi P_B - N_L - U_L + C_L + P_L + A \end{bmatrix}. \quad (E.7)$$

The Jacobian matrix of the equilibrium point  $E8(1, 1, 1)$  is as follows:

$$J_8 = \begin{bmatrix} -R_B + \alpha S_B - R_L + \beta S_L - H & 0 & 0 \\ 0 & P_B - N_B - \alpha S_B - U_B + C_B - B & 0 \\ 0 & 0 & -\phi P_B - N_L - \beta S_L - U_L + C_L + P_L + A \end{bmatrix}. \quad (E.8)$$

## Data Availability

The data used to support the findings of the study are available from the corresponding author upon request.

## Disclosure

The funder has no role in study design, data collection and analysis, publication decisions, or manuscript preparation.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported by the Key Project Fund of Liaoning Provincial Federation Social Science Circles (no. 2021sljdzdk-010), the Social Science Planning Fund of Liaoning Province (L19BGL029), the Key Project Fund of Scientific Research of the Shenyang University of Technology (no. ZDZRGD2020041), and the Humanities and Social Sciences Foundation of Ministry of Education of China (no. 20YJC630195).

## References

- [1] R. Botsman, *What's Mine Is Yours the Rise of Collaborative Consumption*, Harper Business, London, UK, 2011.
- [2] R. Belk, "Why not share rather than own?" *The Annals of the American Academy of Political and Social Science*, vol. 611, no. 1, pp. 126–140, 2007.
- [3] P. A. Albinsson and B. Y. Perera, "Alternative marketplaces in the 21st century: building community through sharing events," *Journal of Consumer Behaviour*, vol. 11, no. 4, pp. 303–315, 2012.
- [4] K. Zhou, L. Y. He, and Y. W. Zhang, "A review of literature on the concept, impacts, and spatial interactions of sharing short-term rental platform," *Progress in Geography*, vol. 39, no. 11, pp. 1934–1943, 2020.
- [5] V. Grybaitė and J. Stankevičienė, "An empirical analysis of factors affecting sharing economy growth," *Oeconomia Copernicana*, vol. 9, no. 4, pp. 635–654, 2018.
- [6] S. Netter, E. Pedersen, and F. Lüdeke-Freund, "Sharing economy revisited: towards a new framework for understanding sharing models," *Journal of Cleaner Production*, vol. 221, pp. 224–233, 2019.
- [7] S. Garcia-Ayllon, "Urban transformations as an indicator of unsustainability in the P2P mass tourism phenomenon: the Airbnb case in Spain through three case studies," *Sustainability*, vol. 10, no. 8, pp. 3–21, 2018.
- [8] M. Ranjbari, G. Morales-Alonso, and R. Carrasco-Gallego, "Conceptualizing the sharing economy through presenting a comprehensive framework," *Sustainability*, vol. 10, no. 7, pp. 2–24, 2018.
- [9] S. Hu, X. Yang, and Q. Wang, "Research review on sharing accommodation at home and abroad," *Tourism Science*, vol. 34, no. 2, pp. 41–57, 2020.
- [10] Y. Yang, K. P.-S. Tan, and X. Li, "Antecedents and consequences of home-sharing stays: evidence from a nationwide household tourism survey," *Tourism Management*, vol. 70, pp. 15–28, 2019.
- [11] F. Long, J. Liu, H. Zhu, and T. Li, "Spatial distribution and influencing factors of homestays in Yangtze River Delta," *Geography Research*, vol. 38, no. 4, pp. 950–960, 2019.
- [12] H. Zhang, L. Lu, D. Zhang, H. Yu, and X. Zhang, "Temporal and spatial distribution characteristics and causes of homestays around Moganshan," *Geography Research*, vol. 38, no. 11, pp. 2695–2715, 2019.
- [13] L. He, "Analysis of the business model of online short rent enterprises - take Piggy short rent as an example," *Modern Business*, vol. 11, no. 9, pp. 44–45, 2016.
- [14] J. Reid, K. Wilson, K. E. Anderson, and C. P. J. Maguire, "Older inpatients' room preference: single versus shared accommodation," *Age and Ageing*, vol. 44, no. 2, pp. 331–333, 2015.
- [15] C. Ling and Z. Zhang, "Research on the development path of 'sharing economy' in China - take online short rent as an example," *Modern Management Science*, vol. 10, pp. 36–38, 2014.
- [16] Y. Wang and L. Yang, "Inspiration of Airbnb business model to China Online short rent Industry," *Modern Business*, no. 3, pp. 175–176, 2017.
- [17] C. Zhao, "Sharing economy: an analysis of the competition structure of China online short rent industry based on five forces model," *Economic Research Guide*, vol. 18, pp. 35–36+182, 2016.
- [18] Y. Xiao, "How to solve 'acclimatization' in online short rent," *Legal Person*, vol. 8, pp. 32–34+96, 2015.
- [19] L. Li and J. Su, "Shared accommodation: the change and influence of the relationship between subject and object," *Tourism Forum*, vol. 12, no. 3, pp. 15–21, 2019.
- [20] S. J. Bae, H. Lee, E.-K. Suh, and K.-S. Suh, "Shared experience in pretrip and experience sharing in posttrip: a survey of Airbnb users," *Information & Management*, vol. 54, no. 6, pp. 714–727, 2017.
- [21] D. Guttentag, "Airbnb: disruptive innovation and the rise of an informal tourism accommodation sector," *Current Issues in Tourism*, vol. 18, no. 12, pp. 1192–1217, 2015.
- [22] C. K. H. Lee and Y. K. Tse, "Improving peer-to-peer accommodation service based on text analytics," *Industrial Management & Data Systems*, vol. 121, no. 2, pp. 209–227, 2021.
- [23] X. Zhang, R. Sui, B. Dan, and Z. Guan, "Shared accommodation: the change and influence of the relationship between

- subject and object,” *Computers & Industrial Engineering*, vol. 155, Article ID 107196, 2021.
- [24] J. Li, A. Moreno, and D. J. Zhang, “Agent behavior in the sharing economy: evidence from Airbnb,” *SSRN Electronic Journal*, vol. 1298, pp. 1–35, 2015.
- [25] L. Kwok and K. L. Xie, “Pricing strategies on Airbnb: are multi-unit hosts revenue pros?” *International Journal of Hospitality Management*, vol. 82, pp. 252–259, 2019.
- [26] C. Gibbs, D. Guttentag, U. Gretzel, J. Morton, and A. Goodwill, “Pricing in the sharing economy: a hedonic pricing model applied to Airbnb listings,” *Journal of Travel & Tourism Marketing*, vol. 35, no. 1, pp. 46–56, 2018.
- [27] Y. Guo, X. Li, and X. Zeng, “Platform competition in the sharing economy: understanding how ride-hailing services influence new car purchases,” *Journal of Management Information Systems*, vol. 36, no. 4, pp. 1043–1070, 2019.
- [28] H. Yu, L. Tian, G. Jiang, and C. Yun, “Sharing economy: theory and research agenda,” *Nankai Business Review*, vol. 21, no. 6, pp. 37–52, 2018.
- [29] D. Guttentag, S. Smith, L. Potwarka, and M. Havitz, “Why tourists choose Airbnb: a motivation-based segmentation study,” *Journal of Travel Research*, vol. 57, no. 3, pp. 342–359, 2018.
- [30] K. K. F. So, H. Oh, and S. Min, “Motivations and constraints of Airbnb consumers: findings from a mixed-methods approach,” *Tourism Management*, vol. 67, pp. 224–236, 2018.
- [31] J. Hamari, M. Sjöklint, and A. Ukkonen, “The sharing economy: why people participate in collaborative consumption,” *Journal of the Association for Information Science and Technology*, vol. 67, no. 9, pp. 2047–2059, 2016.
- [32] M. Mhlmann, “Collaborative consumption: determinants of satisfaction and the likelihood of using a sharing economy option again,” *Journal of Consumer Behaviour*, vol. 14, no. 3, pp. 193–207, 2015.
- [33] S. Konak and Y. Özhasar, “The reasons of foreign tourists choose of Airbnb when they has visited Istanbul,” *International Journal of Contemporary Tourism Research*, vol. 4, no. 1, pp. 79–90, 2020.
- [34] J. Wu, M. Zeng, and K. Xie, “Chinese travelers’ behavioral intentions toward room-sharing platforms,” *International Journal of Contemporary Hospitality Management*, vol. 29, no. 9, pp. 2688–2707, 2017.
- [35] E. Bucher, C. Fieseler, and C. Lutz, “What’s mine is yours (for a nominal fee) - e,” *Computers in Human Behavior*, vol. 62, pp. 316–326, 2016.
- [36] A. Ahani, M. Nilashi, O. Ibrahim, L. Sanzogni, and S. Weaven, “Market segmentation and travel choice prediction in Spa hotels through TripAdvisor’s online reviews,” *International Journal of Hospitality Management*, vol. 80, pp. 52–77, 2019.
- [37] F. Wei, N. Feng, S. Yang, and Q. Zhao, “A conceptual framework of two-stage partner selection in platform-based innovation ecosystems for servitization,” *Journal of Cleaner Production*, vol. 262, Article ID 121431, 2020.
- [38] B. Jiang and L. Tian, “Collaborative consumption: strategic and economic implications of product sharing,” *Management Science*, vol. 64, no. 3, pp. 1171–1188, 2018.
- [39] Z. Shi, R. Cai, and X. Zhu, “Study of business model innovation for intelligent production sharing,” *China Soft Science*, vol. 6, pp. 130–139, 2017.
- [40] K. Lu, J. Zhou, P. Ju, and Y. Xu, “Research on pricing strategy of mobile travel platform based on bilateral market theory,” *Price Theory and Practice*, vol. 11, pp. 150–153, 2016.
- [41] M. ter Huurne, A. Ronteltap, R. Corten, and V. Buskens, “Antecedents of trust in the sharing economy: a systematic review,” *Journal of Consumer Behaviour*, vol. 16, no. 6, pp. 485–498, 2017.
- [42] P. Mikalef, M. Boura, G. Lekakos, and J. Krogstie, “The role of information governance in big data analytics driven innovation,” *Information & Management*, vol. 57, no. 7, Article ID 103361, 2020.
- [43] X. Wu and J. Qiu, “A study of Airbnb listing price determinants: based on data from 36 cities in China,” *Tourism Tribune*, vol. 4, pp. 13–28, 2019.
- [44] C. Wang and H. Chen, “Study on the price of shared short-term rental platform and its influencing factors: based on the data from xiaozhu.com,” *Price: Theory and Practice*, vol. 38, no. 6, pp. 14–17, 2018.
- [45] V. Perez-Sanchez, L. Serrano-Estrada, P. Marti, and R.-T. Mora-Garcia, “The what, where, and why of Airbnb price determinants,” *Sustainability*, vol. 10, no. 12, pp. 4596–4627, 2018.
- [46] Y. Li, S. Wang, Y. Ma, Q. Pan, and E. Cambria, “Popularity prediction on vacation rental websites,” *Neurocomputing*, vol. 412, pp. 372–380, 2020.
- [47] A. Lawani, M. R. Reed, T. Mark, and Y. Zheng, “Reviews and price on online platforms: evidence from sentiment analysis of Airbnb reviews in Boston,” *Regional Science and Urban Economics*, vol. 75, pp. 22–34, 2019.
- [48] D. Wang and J. L. Nicolau, “Price determinants of sharing economy based accommodation rental: a study of listings from 33 cities on Airbnb.com,” *International Journal of Hospitality Management*, vol. 62, pp. 120–131, 2017.
- [49] E. Ert, A. Fleischer, and N. Magen, “Trust and reputation in the sharing economy: the role of personal photos in Airbnb,” *Tourism Management*, vol. 55, pp. 62–73, 2016.
- [50] Q. Xu, C. Wu, and Q. Chen, “Dynamic optimal allocation strategy of idle resources in supply chain under shared platform,” *Operations Research and Management Science*, vol. 30, no. 9, pp. 86–92, 2021.
- [51] G. L. Lunardi, A. C. G. Maçada, J. L. Becker, and W. Van Grembergen, “Antecedents of IT governance effectiveness: an empirical examination in Brazilian firms,” *Journal of Information Systems*, vol. 31, no. 1, pp. 41–57, 2017.
- [52] Z. Alreemy, V. Chang, R. Walters, and G. Wills, “Critical success factors (CSFs) for information technology governance (ITG),” *International Journal of Information Management*, vol. 36, no. 6, pp. 907–916, 2016.
- [53] L. Piscicelli, T. Cooper, and T. Fisher, “The role of values in collaborative consumption: insights from a product-service system for lending and borrowing in the UK,” *Journal of Cleaner Production*, vol. 97, no. 15, pp. 21–29, 2015.

## Research Article

# Lightweight YOLOv4 with Multiple Receptive Fields for Detection of Pulmonary Tuberculosis

Zhitao Guo , Jiahao Wang, Jinghua Wang, and Jinli Yuan 

College of Electronic Information Engineering, Hebei University of Technology, Tianjin 300401, China

Correspondence should be addressed to Jinli Yuan; jinli\_yuan@hebut.edu.cn

Received 1 February 2022; Revised 17 February 2022; Accepted 21 February 2022; Published 31 March 2022

Academic Editor: Daqing Gong

Copyright © 2022 Zhitao Guo et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The characteristics of pulmonary *tuberculosis* are complex, and the cost of manual screening is high. The detection model based on convolutional neural network is an essential method for assisted diagnosis with artificial intelligence. However, it also has the disadvantages of complex structure and a large number of parameters, and the detection accuracy needs to be further improved. Therefore, an improved lightweight YOLOv4 pulmonary *tuberculosis* detection model named MIP-MY is proposed. Firstly, over 300 actual cases are selected to make a common dataset by professional physicians, which is used to evaluate the performance of the model. Subsequently, by introducing the inverted residual channel attention and the pyramid pooling module, a new structure of MIP is created and used as the backbone extractor of MIP-MY, which could further decrease the number of parameters and fuse context information. Then the multiple receptive field module is added after the three effective feature layers of the backbone extractor, which effectively enhances the information extraction ability of the deep feature layer and reduces the miss detection rate of small pulmonary *tuberculosis* lesions. Finally, the pulmonary *tuberculosis* detection model MIP-MY with lightweight and multiple receptive field characteristics is constructed by combining each improved modules with multiscale structure. Compared to the original YOLOv4, the model parameters of MIP-MY is reduced by 47%, while the mAP value is raised to 95.32% and the miss detection rate is decreased to 6%. It is verified that the model can effectively assist radiologists in the diagnosis of pulmonary *tuberculosis*.

## 1. Introduction

Pulmonary *tuberculosis* has become a global public health emergency, in which people are infected with a chronic infectious disease caused by *Mycobacterium tuberculosis* [1]. According to the World Health Organization (WHO) survey, pulmonary *tuberculosis* is one of the leading causes of death from infectious diseases around the world. It is estimated that about 64% of 10 million pulmonary *tuberculosis* cases are detected and treated each year [2]. Improving the awareness rate and early detection rate of pulmonary *tuberculosis* plays a vital role in the treatment of the disease, as well as the prevention of the spread of the disease [3]. Computed tomography (CT) is one of the auxiliary imaging diagnostic methods for *tuberculosis* screening, having a lower missed detection and false detection rate than chest radiographs. It is a more efficient choice to adopt CT

technology to identify substantial chest lesions and detect the severity of lung diseases in patients with *tuberculosis* [4].

With the development of artificial intelligence, some scholars have gradually begin to integrate the deep learning image processing algorithm of deep learning with CT technology to achieve a more accurate diagnosis and detection of lung diseases. Many deep learning models of computer-aided diagnosis have been built based on the deep convolution neural network (DCNN). Gao et al. [5] combined with CT technology proposed a high-precision classification model of five types of pulmonary *tuberculosis* based on CNN and support vector machine. Ma et al. [6] proposed an automatic detection model of active pulmonary *tuberculosis* based on U-Net [7], which can detect the location of lesions more accurately. Liu et al. [8] built simulated and real data sets based on CT images of lung cancer lung nodules and realized automatic detection of lesions

through an improved single-stage target detection network YOLOv3 [9]. Yang et al. [10] took the two-stage target detection network Faster-RCNN as the main body, which improved the detection accuracy of pulmonary *tuberculosis*. However, its parameter scale was large and was not conducive to generalization.

Considering that the current detection model for pulmonary CT disease has the problem of excessive parameters and does not use effective lightweight methods, it may be a feasible strategy to learn from the lightweight methods in other detection fields. In reference [11], Ye et al. [11] replaced the backbone network of YOLOv4 [12] through MobileNetv3 [13] and realized the lightweight of the model. Although Mobilenetv3 reduces the amount of parameters, if it is directly used in the detection of pulmonary *tuberculosis* CT, it may lack the mining of deep feature information of the image, which is not enough to solve the problem of missed detection rate of small targets by YOLOv4, and the detection accuracy needs to be further improved. Reference [14] achieves lightweight by reducing the number of convolution layers in YOLOv5 [15], which obviously improves the detection speed. However, because the lightweight YOLOv5 focuses on the detection of large targets, the detection performance of small targets is still slightly insufficient. Small lesions on pulmonary *tuberculosis* CT cannot be ignored. The structure and performance of YOLOv5 are less different from those of YOLOv4. Therefore, improving on YOLOv4 is an effective strategy. CT images of pulmonary *tuberculosis* have mainly have typical features such as cavity [16] and tree-in-bud pattern [17]. However, open data sets are uneven and there is no uniform format, so it is essential to design a set of standard datasets. Additionally, the existing *tuberculosis* detection model has a large number of parameters and consumes a lot of computing resources, which makes it difficult to be applied to medical institutions with a large number of high-performance equipment. Therefore, it is an important goal to realize the lightweight of the model. At the same time, an ordinary lightweight detection model cannot obtain the deeper semantic information of CT images, resulting in a high rate of missed detection. Therefore, ensuring high detection accuracy and efficiency of *tuberculosis* simultaneously is one of the main objectives of this study, which is of great significance for *tuberculosis* screening and diagnosis.

The major contributions in our study can be summarized as follows.

- (1) At present, some existing models lack the ability to distinguish small objects, so that the miss detection rate stays in a high position without going down. For this reason, a new module of the multiple receptive field block (MRFB) with dilated convolution is designed in this paper. This module expands the acceptance range of the pulmonary *tuberculosis* detection model. It enhances the ability of feature extraction for various sizes of lesions to prevent some small size lesions that are detrimental to identify from being omitted.

- (2) Location, classification, and the confidence score of pulmonary *tuberculosis* lesions are equally important factors affecting the detection capability of the model. Taking these three factors as optimization objects can further improve the detection accuracy of the model. Consequently, an integrated loss function is designed which can train these three factors simultaneously. In the iterative training process of the model, the actual location of the lesions will be continuously updated continuously by the integrating loss function, the classification of the lesions will be determined, and the reliability of all lesions will be optimized simultaneously.
- (3) Lightweight convolution layers are constructed to reduce the number of model parameters, and a new backbone extractor MIP is designed. Its internal inverted residual channel attention module and pyramidal pooling layer can assist the model in extracting the deeper semantic information of pulmonary *tuberculosis* lesions. It realizes lightweight in structure and avoids overfitting to a certain extent, enhances the relationship between the different sizes of regions on feature maps, and completes the fusion of multiscale feature information.

The remainder of this paper is organized as follows. In Section 1, a standard data set of pulmonary *tuberculosis* is built, which contains CT slices of 300 patients with pulmonary *tuberculosis*. And a lightweight method is also introduced that is used in the pulmonary *tuberculosis* detection model. In Section 2, a pulmonary *tuberculosis* detection model MIP-MY is designed based on an improved YOLOv4 algorithm. In section 3, we carry out a quantitative analysis and score comparison on the results of networks. Finally, we discuss some related issues and make conclusions in Section 4.

## 2. Dataset and Lightweight Method

Currently, the public data set on CT of pulmonary *tuberculosis* CT is scarce, with incomplete image data, and the format is not unified. Therefore, it is imperative to design a standard data set for the research of this study. The lightweight detection model is one of the targets to be achieved in this paper. While completing the lightweight, we must ensure a higher accuracy, which is also the difficulty in designing the pulmonary *tuberculosis* detection model.

**2.1. Creation of Dataset.** The experimental data for this study is the actual data set provided by the Imaging Department of Beijing Chest Hospital, China. A total of 300 CT cases of pulmonary *tuberculosis* are collected to form the experimental data set. Each CT slice is segmented according to the thickness standard of 1.25 mm, of the which the CT sections of pulmonary *tuberculosis* cavity and the tree-in-bud pattern accounted for 50%, respectively. The cavity diameter in the data set ranges from 10 mm to 126.4 mm, and the tree-in-bud pattern with uniform density is selected as the sample. The areas from  $16 \times 16$  to  $128 \times 128$  on the CT slices are



defined as a cluster of tree-in-bud pattern lesions, respectively. The distribution maps of the size of the lesions size in the dataset are shown in Figure 1. As shown from Figures 1(a) and 1(b), there are more small cavities with a diameter less than or equal to 30 mm, accounting for about 33% of the total. And more samples of tree-in-bud patterns in clusters with the size from  $32 \times 32$  to  $64 \times 64$  account for about 40% of the total. The standard for the above data set was established by three clinical doctors with more than five years of experience at Beijing Chest Hospital after discussion at the meeting, and the Labellmg tool [18] tool was used to outline and calibrate the lesion of pulmonary *tuberculosis*, providing a guarantee of the consistency and validity of the dataset.

The unified preprocessing of each CT image is carried out, and its format is converted from DICOM to PNG, which is more convenient for processing. The resolution is adjusted to  $512 \times 512$ , and the number of coding bits is 24. Additionally, to reduce interference to the detection model, background information is filtered, including bed board and clothing outside the outline of the lung.

The data set contains a total of 3764 CT slices of cavity and tree-in-bud pattern, from which 70% are randomly selected as the train set, the remaining 10% as test set 1, and the last 20% as test set 2. Test Set 1 is responsible for subsequent ablation experiments, while test set 2 is responsible for evaluating the diagnostic level of the pulmonary *tuberculosis* detection model.

Data enhancement enriches the diversity of the data set. Specifically, each CT image of the train set has a 40% chance of scaling and flipping horizontally and randomly distorting the input image with an aspect ratio of 0.8 to 2.0. These measures double the number of samples in the train set. This study has signed a patient information confidentiality agreement with the Imaging Department of Beijing Chest Hospital to filter out patients' sensitive information by technical means, and the right to use the dataset is merely effective in this study.

**2.2. Lightweight Method.** At present, the detection model based on the convolutional neural network has some problems such as a large number of parameters and long training time, which requires high hardware computing capacity. Therefore, this paper will adopt a lightweight method to modify the model to make it suitable for general hardware devices.

MobileNetv3 is a lightweight neural network commonly used in image processing. It is based mainly on the principle of separable depthwise convolution [19] to achieve the lightweight target.

Compared with traditional convolution, depthwise separable convolution has fewer parameters. The Depthwise Separable Convolution traverses only one by one corresponding to the input channel and then expands the number of output channels by  $1 \times 1$  Pointwise Convolution. The principle of Depthwise Separable Convolution is shown in Figure 2.

The number of parameters required for traditional convolution and Depthwise Separable Convolution is set to

$N_1$  and  $N_2$  respectively, and their calculation formulas are shown in formulas (1) and (2).

$$N_1 = i_h \cdot i_w \cdot c_{in} \cdot c_{out} \cdot k^2, \quad (1)$$

$$N_2 = i_h \cdot i_w \cdot c_{in} \cdot (k^2 + c_{out}), \quad (2)$$

where  $i_h$  and  $i_w$  are the height and width of the input tensor,  $k$  is the size of the selected convolution kernel, and  $c_{in}$  and  $c_{out}$  respectively represent the number of input and output channels. Furthermore, as shown in formula (3)

$$\frac{N_1}{N_2} = \frac{k^2 \cdot c_{out}}{k^2 + c_{out}}. \quad (3)$$

In practical application,  $N_2$  is much less than  $N_1$ , which means Depthwise Separable Convolution can reduce computational overhead to a greater extent.

The model in reference [11] replaced YOLOv4's backbone extractor CSPDarkNet53 [20] by MobileNetv3, which reduces the number of parameters to a great extent. However, it is difficult to achieve high accuracy if this model is directly used to detect pulmonary *tuberculosis*. If the lightweight of the model is improved at the expense of detection accuracy, that will be contrary to the main target of our study. Therefore, this article will ameliorate the MobileNetv3 and YOLOv4 to reduce the parameters of the pulmonary *tuberculosis* detection model, improve detection efficiency as well as the accuracy.

### 3. Design of Pulmonary Tuberculosis Detection Model

The structure of the pulmonary *tuberculosis* detection model refers to the YOLOv4 method, and the structure of YOLOv4 is shown in Figure 3. As can be seen from Figure 3, the original YOLOv4 enhances the learning feature information ability by virtue of the CSPDarkNet53 as the backbone part, which reduces model overfitting through the combination of traditional convolution and residuals. The neck part is spliced with SPP module (spatial pyramid pooling) [21] and PAN module (path aggregation network) [22] to complete the multiscale feature information fusion of different regions. Finally, the multiscale feature information is collected by using the three Heads in the prediction part to generate the final prediction bounding box. The original YOLOv4 focuses on efficiency and accuracy, but has not yet achieved satisfactory results in the field of pulmonary *tuberculosis* detection.

Different from the original YOLOv4, the developed pulmonary *tuberculosis* detection model named MIP-MY comprises three parts: a backbone extractor, enhanced feature extractor, and a bounding box generator. A new backbone extractor MIP is designed for MIP-MY, which is the lightweight module of the detection model. MIP uses the newly designed IRCA (Inverted Residual Channel Attention) module and pyramid pooling module to achieve lightweight. MIP can extract abundant contextual information and make preliminary sampling for pulmonary

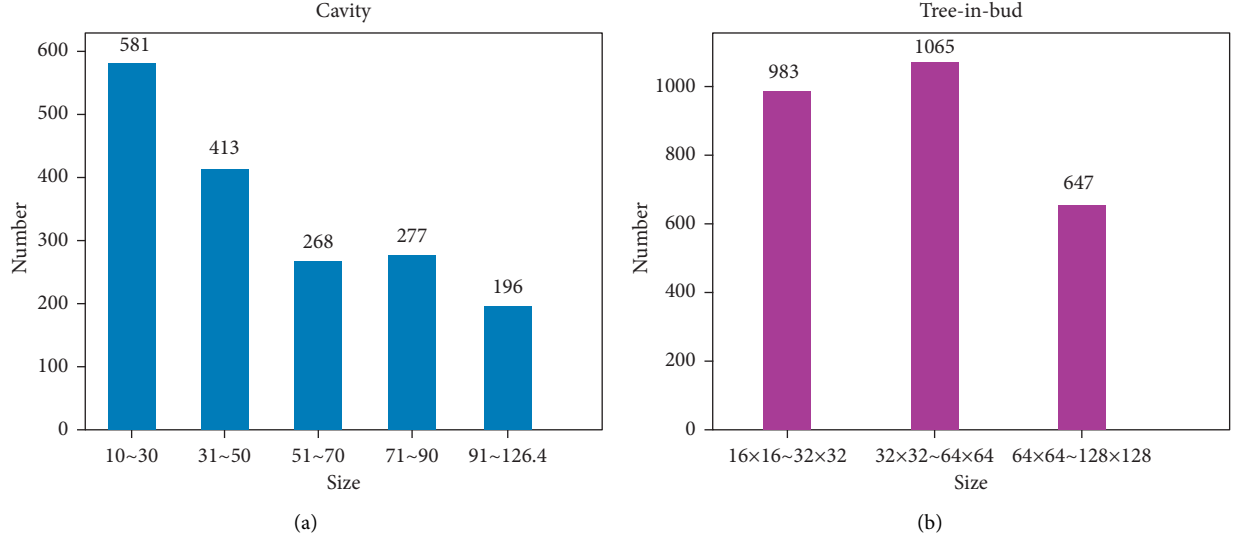


FIGURE 1: Distribution maps of the size of the lesions size. (a) Distribution maps of the cavity. (b) Distribution maps of the tree-in-bud pattern.

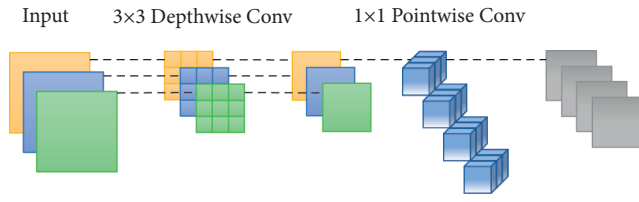


FIGURE 2: Principle of depthwise separable convolution.

*tuberculosis*. Then, in order to capture lesions of different sizes, we designed MRFB (multiple accept field block) module to construct the enhanced feature extractor, which can have a wider acceptance field to reduce the missed detection of small lesions. According to the momentous information (lesion location, classification, and confidence score) output by the bounding box generator, a comprehensive loss function is designed to regression the lesion location and optimize the confidence score.

**3.1. Backbone Extractor MIP.** The function of the backbone feature extractor MIP is to analyze the CT images of pulmonary *tuberculosis* and extract the image features of pulmonary *tuberculosis* lesions. In this paper, according to the inverted residuals structure of MobileNetv3, a new IRCA (Inverted Residuals Channel Attention) module is constructed using channel attention mechanism. By virtue of the inverted residuals structure, this module can build a deeper network without a gradient explosion. An introduced channel attention mechanism can capture the corresponding context information on each IRCA module and improve the feature extraction ability of the network. The structure of IRCA is shown in Figure 4. As a significant part of MIP, IRCA module constructs seven effective feature layers for MIP to extract the lesions information of pulmonary *tuberculosis*. The MIP connects a pyramid pool module in the last effective feature

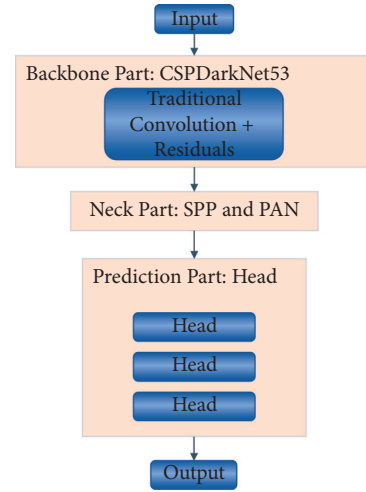


FIGURE 3: The structure of YOLOv4.

layer to obtain multiple scale feature information. The structure of MIP and pyramid pooling module is shown in Figure 5.

Before completing the  $1 \times 1$  Pointwise Convolution operation, the IRCA module carries out Global-Average pooling to obtain the feature graph with a smaller size, and calculate the weight of the feature map through Relu6 and Hard-Swish activation function. Finally, the weighted multiplication is performed through channel attention. The calculation formula of Global-Average pooling is shown in formula (4).

$$y_{\text{avg}} = \frac{\sum_{i=0}^n \sum_{j=0}^n x_{i,j}}{n^2}, \quad (4)$$

where  $x_{i,j}$  is the pixel on the 2D slices of feature maps and  $y_{\text{avg}}$  is the pixel mean of the Global-Average pooling. The calculation formula of Relu6 and Hard-Swish activation function is shown in formulas (5) and (6).

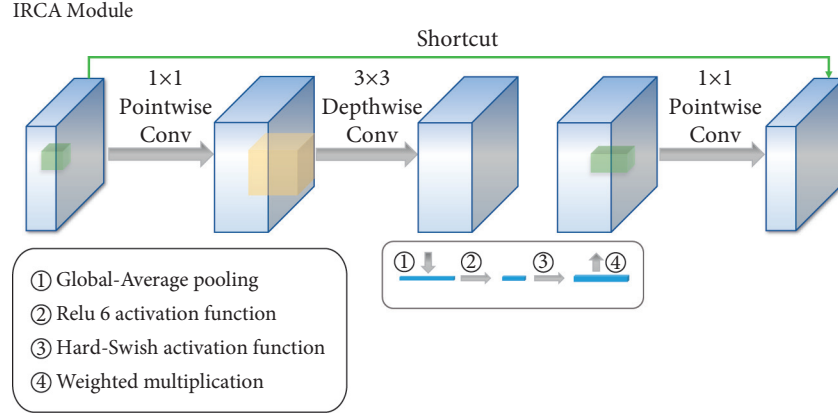


FIGURE 4: The structure of the IRCA module.

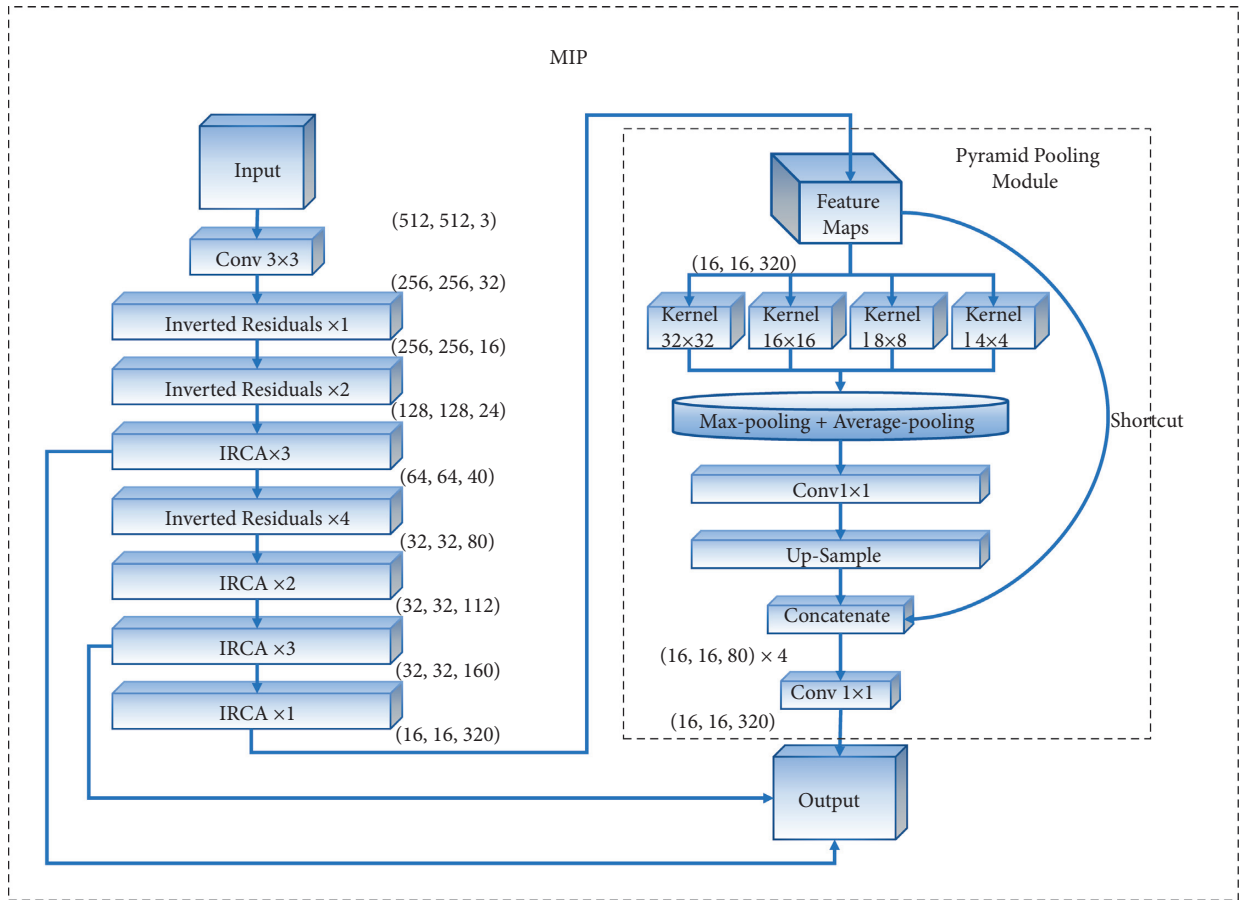


FIGURE 5: The structure of MIP and pyramid pooling module.

$$f_1(y_{\text{avg}}) = \min(\max(0, y_{\text{avg}}), 6), \quad (5)$$

$$f_2(y_{\text{avg}}) = y_{\text{avg}} \cdot \frac{f_1(y_{\text{avg}} + 3)}{6}, \quad (6)$$

where  $f_1(y_{\text{avg}})$  is the activation result of ReLU6, and the upper limit of this value is 6.  $f_2(y_{\text{avg}})$  is the activation result of Hard-Swish. It is found that the accuracy of a neural network can be enhanced by using the hard-swish activation function in deeper convolution layers.

$$f_3 = X \cdot f_2(f_1(y_{\text{avg}})), \quad (7)$$

where the output  $X$  of the ‘‘Depthwise Separable Convolution’’ in formula (7) is a characteristic graph composed of  $n^2$  pixels, and  $f_3$  is the weighted calculation result of the whole channel attention.

In order to enable the backbone extractor to acquire information about the lesions information of different scales, this design draws lessons from the structure of PSPNet (Pyramid Scene Parsing Network) [23] and designs a

pyramid pooling module as the integrated pooling layer of the backbone extractor (max pooling + average pooling), which combines four pooled kernels of various scales ( $32 \times 32$ ,  $16 \times 16$ ,  $8 \times 8$  and  $4 \times 4$ ) to further analyze the location, size and other information about the lesions.

In summary, this design is improved by using the IRCA module and pyramid pooling module to compose the structure of a new backbone extractor MIP.

**3.2. Enhanced Feature Extractor Dominated by MRFB.** The last three effective feature layers of the MIP backbone extractor MIP will export feature maps with the resolution of the original image of 1 to 8, 1 to 16, and 1 to 32, respectively, providing pixel information of different sizes of lesions in pulmonary *tuberculosis* for subsequently enhanced feature extractor.

The initially enhanced feature extractor of YOLOv4 is mainly composed of SPP module and PAN module. SPP will take the result of the last feature layer of the backbone extractor as input and divide the feature map into three subregions of different sizes ( $8 \times 8$ ,  $4 \times 4$ , and  $2 \times 2$ ) for maximum pooling, collecting multiscale eigenvalues. The PAN is a reciprocating structure responsible for collecting the output of three effective feature layers of the backbone extractor to build a feature pyramid. Each effective feature layer integrates the sampling information of the other two layers to complete multiscale feature lesions.

Although the initially enhanced feature extractor uses a multiscale feature fusion method to collect image information, the utilization of multireceptive field information is far from enough, and it is easy to ignore the feature expression of small targets. Therefore, in this section, combining the ideas of expansion convolution and receptive field amplification [24], the MRFB module is designed to optimize the enhanced feature extractor of YOLOv4. The structure of MRFB is shown in Figure 6.

Combined with the structure of the MRFB module, it can see that the MRFB module adopts parallel expansion convolution, and expands and samples the receptive field of the feature map through the collocation of three expansion rates (dilated rate takes 1, 3 and 5) and three convolution kernels (size takes  $1 \times 1$ ,  $3 \times 3$ , and  $5 \times 5$ ), and adds residual shortcut to prevent the loss of feature information, and finally splices the sampling results of each convolution to fuse the information of multiple receptive fields. As shown in Figure 6, the MRFB module, as the probe of PAN, is connected with the backbone feature extractor MIP, which makes up for the deficiency of PAN feature map information collection.

**3.3. Bounding Box Generator and Decoder.** The function of the bounding box generator is to achieve regression of the information from the pulmonary *tuberculosis* lesion (lesion size, class, and location), which is composed of three decoupling heads and a decoder. Each head generates three bounding boxes of different sizes on the multiscale feature map sampled by PAN to surround the pulmonary *tuberculosis* lesion. The parameters of the bounding box include the coordinates of the center point coordinates, size, class of lesion, and confidence

score. However, this parameter information cannot directly reflect the position of the final bounding box in the picture. The information of the bounding box needs to be further decoded by the decoder.

**3.3.1. Decoding of Bounding Box.** The bounding box is decoded with Anchor. Anchor is a predefined bounding box on each feature point of the input image. The activation function for its decoding is shown in formula (8).

$$\sigma(t) = \frac{1}{1 + e^{-t}}, \quad (8)$$

where  $t$  represents the confidence score and category probability of the bounding box, and these two kinds of parameters are mapped to the range of  $[0,1]$  by function  $\sigma(t)$ . The decoding definition of the center point coordinates  $(C_x, C_y)$  of the bounding box is shown in formulas (9) and (10).

$$C_x = \sigma(x_{\text{offset}}) + A_x, \quad (9)$$

$$C_y = \sigma(y_{\text{offset}}) + A_y, \quad (10)$$

where the coordinate offset ( $x_{\text{offset}}$  and  $y_{\text{offset}}$ ) from the center point of the prediction box is the coordinate offset and the coordinate offset is relative to the center point of the prediction box.  $(A_x, A_y)$  is the center point coordinate of the Anchor, and the center point coordinates  $(C_x, C_y)$  of the bounding box is obtained after decoding.

The decoding formulas for the height and width of the bounding box are shown in formulas (11) and (12).

$$H = e^h \cdot A_H, \quad (11)$$

$$W = e^w \cdot A_W, \quad (12)$$

where the  $h$  and  $w$  are the height and width of the bounding box before decoding,  $A_H$  and  $A_W$  are the height and width of Anchor, and the height  $H$  and width  $W$  of the decoded bounding box is calculated, respectively.

After decoding, the pulmonary *tuberculosis* lesions at the same position in the image will be surrounded by a large number of bounding boxes, so it is necessary to filter the redundant bounding boxes according to the Intersection of Union threshold (IOU) (usually set to 0.5), and then filter out these bounding boxes with the highest confidence score through the nonmaximum suppression algorithm [25]. The mathematical definition of the IOU is shown in formula (13).

$$\text{IOU} = \frac{|b \cap b^{gt}|}{|b \cup b^{gt}|}, \quad (13)$$

where the denominator represents the intersection of the area of the bounding box  $b$  and the real box  $b^{gt}$ , and the numerator represents the union. IOU reflects the similarity between the predicted results of the detection model and the ground truth.

**3.3.2. Integrated Loss Function.** The integrated loss function of the pulmonary *tuberculosis* detection network is mainly composed of regression loss  $L_{\text{CIou}}$ , class loss  $L_{\text{Class}}$  and confidence loss  $L_{\text{Conf}}$ .

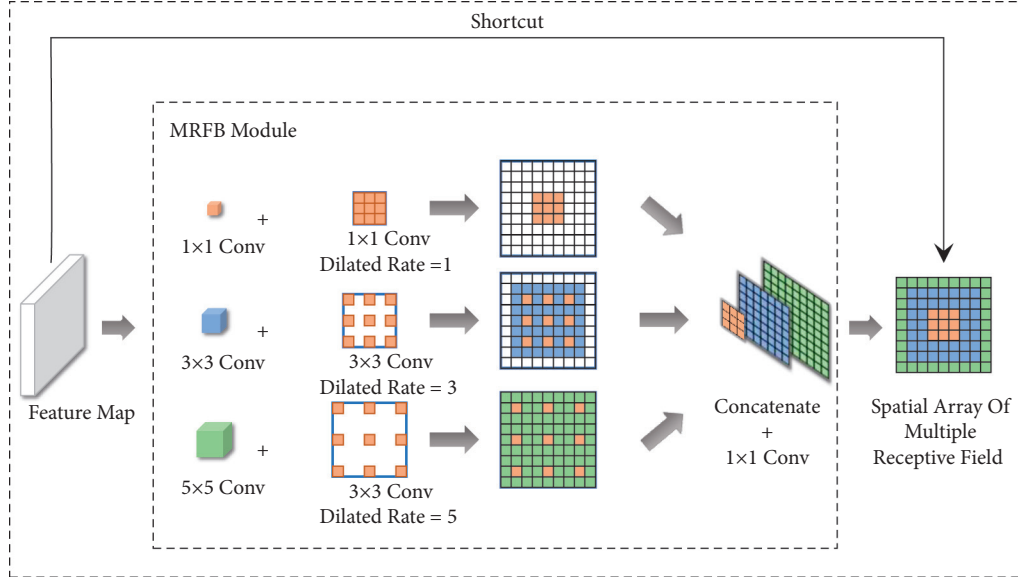


FIGURE 6: The structure of MRFB module.

The general loss function of the detection model is shown in formula (14).

$$L = L_{\text{CIou}} + L_{\text{Confidence}} + L_{\text{Class}} \quad (14)$$

The mathematical expression of regression loss  $L_{\text{CIou}}$  is shown in formula (15).

$$\begin{cases} v_{i,j} = \frac{4}{\pi^2} \left( \tan^{-1} \frac{W_{i,j}^{gt}}{H_{i,j}^{gt}} - \tan^{-1} \frac{W_{i,j}}{H_{i,j}} \right)^2, \\ \alpha_{i,j} = \frac{v_{i,j}}{1 - \text{IOU}_{i,j} + v_{i,j}}, \\ L_{\text{CIou}} = \sum_i^{S^2} \sum_j^N \left[ 1 - \text{IOU}_{i,j} + \frac{\rho(b_{i,j}, b_{i,j}^{gt})}{d_{i,j}^2} + \alpha_{i,j} v_{i,j} \right], \end{cases} \quad (15)$$

where the subscript  $(i, j)$  shows the serial number of the  $j$  bounding box on the pixel point in the feature map,  $S$  is the resolution size of the feature maps and  $i$  is the number of predicted bounding boxes on every feature map.  $W_{i,j}^{gt}$  and  $H_{i,j}^{gt}$  are the width and height of the corresponding ground truth,  $\alpha_{i,j} v_{i,j}$  is the penalty factor of  $L_{\text{CIou}}$ , and  $\rho(b_{i,j}, b_{i,j}^{gt})$  is the Euclidean distance between the center point of the predicted bounding box and the ground truth, and  $d_{i,j}$  represents the diagonal distance of the smallest enclosed area that can contain both the predicted bounding box and the ground truth. The definition of class loss is shown in formula (16).

$$L_{\text{Confidence}} = \sum_i^{S^2} \sum_j^N O_{i,j} [-\log(E_{i,j})] + \lambda \sum_i^{S^2} \sum_j^N (1 - O_{i,j}) [-\log(1 - E_{i,j})], \quad (16)$$

where the  $O_{i,j}$  is a binary number to judge whether the pulmonary *tuberculosis* focus is in the predicted bounding box, and  $E_{i,j}$  represents the confidence of every pulmonary

*tuberculosis* focus.  $\lambda$  sets as 1 when the IOU is greater than the threshold, otherwise takes 0. The class loss  $L_{\text{Class}}$  is defined in formula (17).

$$L_{\text{Class}} = \sum_i^{S^2} \sum_j^N O_{i,j} \sum_{c \in \text{Classes}}^M [P_i(c) \log(P_i^{gt}(c)) + (1 - P_i(c)) \log(1 - P_i^{gt}(c))], \quad (17)$$

where  $M$  is the number of classes,  $P_i(c)$  is the class score of the detection network decision, and  $P_i^{gt}(c)$  represents class score of the ground truth.

By improving the structure of the pulmonary *tuberculosis* detection model and the selection of the loss function, the *tuberculosis* detection model developed in this paper is named MIP-MY (MobileNet with Inverted residuals and Pyramid pooling - Multiple receptive fields of YOLO). The model creates a lightweight trunk extractor MIP based on the basis of YOLOv4 and uses the MRFB module to replace the ordinary convolution layer that enhances the partial redundancy of the feature extractor, which only takes up less memory to enhance the ability to obtain multiple receptive field features. Finally, the integrated loss function is aimed at completing the regression and classification of pulmonary *tuberculosis* lesion information at the model training.

The overall structure of the improved pulmonary *tuberculosis* detection model MIP-MY is shown in Figure 7.

**3.4. Evaluation Method.** The evaluation of the pulmonary *tuberculosis* detection model is mainly determined by the number of parameters, detection time of per CT, Precision, Recall, miss detection rate, and mean Average Precision (mAP) [26]. The Precision  $F_{Pr}$  is as shown in formula (18).

$$F_{Pr} = \frac{TP}{TP + FP}, \quad (18)$$

where TP is the true positive of the sample, FP is the false positive, and the sum of them is the prediction result of the model. The recall rate  $F_{Re}$  is shown in formula (19).

$$F_{Re} = \frac{TP}{TP + FN}, \quad (19)$$

where the FN is false negative, and the sum of FP and FN is the total amount of the real box. In this study, the miss detection rate is mapped from the average miss detection rate of pulmonary *tuberculosis* CT in the test set 2 to the logarithmic space, and its mathematical definition is as shown in formula (20).

$$\left\{ \begin{array}{l} mr_i = 1 - \frac{TP_i}{TP_i + FP_i}, \\ F_{Lamr} = e^{-\frac{1}{Q} \sum_{i=1}^Q \log(\max(mr_i, \varepsilon))}, \end{array} \right. \quad (20)$$

where  $mr_i$  represents the miss rate of a single CT slice and  $Q$  represents the total number of CT slices. Besides,  $\log(0)$  has no mathematical meaning, a smaller value  $\varepsilon$  is set to prevent the independent variable of the logarithmic function from being zero.

The IOU by a threshold of 0.5 is set to determine the TP and FP of the sample, and the P-R curve of the detection model is constructed with Precision and Recall as horizontal and vertical coordinates, respectively. The mAP is equal to

the mean area under the P-R curve of all categories. The mathematical definition of the mAP is shown in formula (21), where  $n$  is the total number of categories.

$$mAP = \frac{\sum_0^n \int E_{pr} \cdot E_{red}(E_{re})}{n}. \quad (21)$$

## 4. Experimental Setting and Analysis

The operating system for this experiment is Windows 10, and the processors are Intel Core i7 and RTX 2060. Take PyCharm as the integrated development environment and utilize the deep learning framework of Keras based on Python3.8. To adapt to model training, the initial learning rate is set at  $1.0 \times 10^{-4}$ . Adam optimizer and 1000 iterative training are also selected. Finally, the cosine annealing algorithm is used so that the pulmonary *tuberculosis* detection model can adaptively adjust the learning rate according to the number of iterations.

**4.1. Ablation Experiment.** For the sake of investigating the contribution of the integrated loss function and different components (MIP and MRFB) to improve the accuracy of detection model MIP-MY, we selected test set 1 to carry out two groups of ablation experiments.

The first group verifies the influence of different sub-losses (regression loss  $L_{CIou}$ , class loss  $L_{Class}$  and confidence loss  $L_{Conf}$ ) in the integrated loss function on the detection performance of the model, and then compares them with the traditional cross-entropy loss.

The quantitative comparison of performance on each loss function is shown in Table 1. Any single sub-loss can not optimize the detection accuracy of the model, and it is accompanied by a huge miss detection rate. Similarly, the combination of any two sub-losses has no obvious improvement on the detection accuracy and miss detection rate. The experimental results show that the coordinate location, class, and confidence score of pulmonary *tuberculosis* lesions are indispensable factors and only by combining these three seed losses can we effectively improve the overall performance of MIP-MY. Compared with traditional loss of the cross-entropy loss function, integrated loss can obtain better training results.

The second group verifies the effectiveness of the improved components in this paper, that is, we observe the effects of the addition and deletion of MIP module and MRFB module on the mAP, the number of parameters, miss detection rate and detection time. And compared with the lightweight model in reference [11]. To facilitate differentiation, we named the lightweight model introduced by Model 1 in reference [11] as MobileNetv3-YOLOv4, the model that uses only MIP is called Model 2 and the model that uses only MRFB is called Model 3. MIP-MY integrates all the improved modules.

The results of the ablation experiment are shown in Table 2. From these data, it can be inferred that, compared with Model 1, the number of parameters of Model 2 is only increased by 0.93 M, but the mAP is increased by 2.64%, and



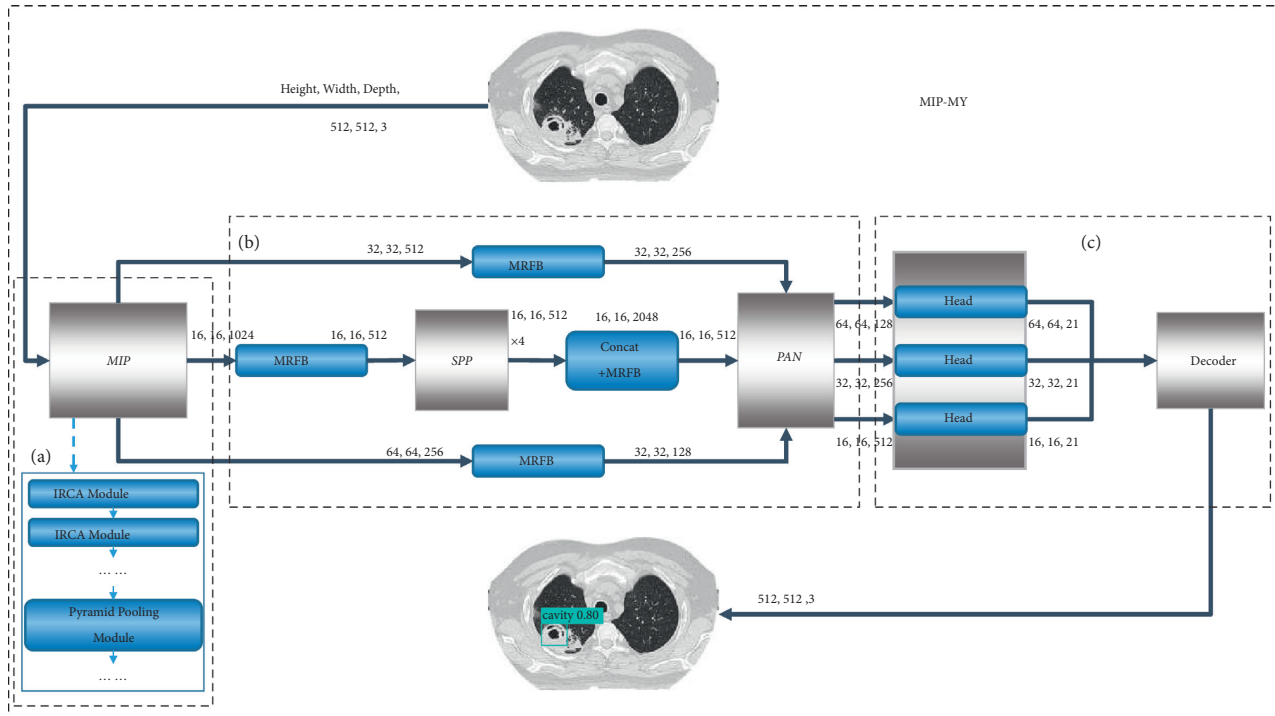


FIGURE 7: The overall structure of MIP-MY (a) Backbone extractor MIP; (b) MRFB enhanced feature extractor with MRFB; (c) boundary box generator and decoder.

TABLE 1: The quantitative comparison of performance on each loss function.

Method	mAP (%)	Miss detection rate (%)
$L_{CIUO}$	79.47	20
$L_{Class}$	78.81	20
$L_{Conf}$	76.15	21
$L_{CIUO} + L_{Class}$	87.41	14
$L_{CIUO} + L_{Conf}$	85.06	17
$L_{Class} + L_{Conf}$	84.29	18
Cross-entropy loss	90.36	11
Integrated loss	95.59	6

TABLE 2: The influence of each improved module on the pulmonary *tuberculosis* detection model.

Model	mAP (%)	Parameters (M)	Miss detection rate (%)	Detection time (s)
Model 1	85.86	38.64	16	6.95
Model 2	88.50	39.57	13	7.87
Model 3	91.92	58.29	10	8.86
MIP-MY	95.59	33.91	6	5.88

the miss detection rate is reduced by 3%. Model 3 only uses MRFB to improve its mAP to 91.92%, but because it does not adopt the lightweight module MIP, the number of parameters is still high and with lower detection efficiency. MIP-MY combines the advantages of two improved modules to make the structure lightweight with the expansion of the receptive field feature mapping. Its mAP jumps to 95.59%, 9.73% higher than model 1, and the miss detection rate decreases by 10%. Meanwhile, the number of parameters decreases by 4.73 M, and the detection time of 1.07 s is shortened.

**4.2. Comparative Experiment of Mainstream Models.** Generally, the *tuberculosis* detection model is highly efficient, highly precise and low memory consumption and easy to apply to clinical diagnosis. To further verify the reliability of the MIP-MY model, we use the test set 2 to compare with several mainstream target detection models. The test set 2 contains 347 cavity samples and 539 tree-in-bud pattern samples. These CT samples are used to evaluate the performance of different detection models.

The P-R curve of the cavity of pulmonary *tuberculosis* cavity and the tree-in-bud pattern detected by different

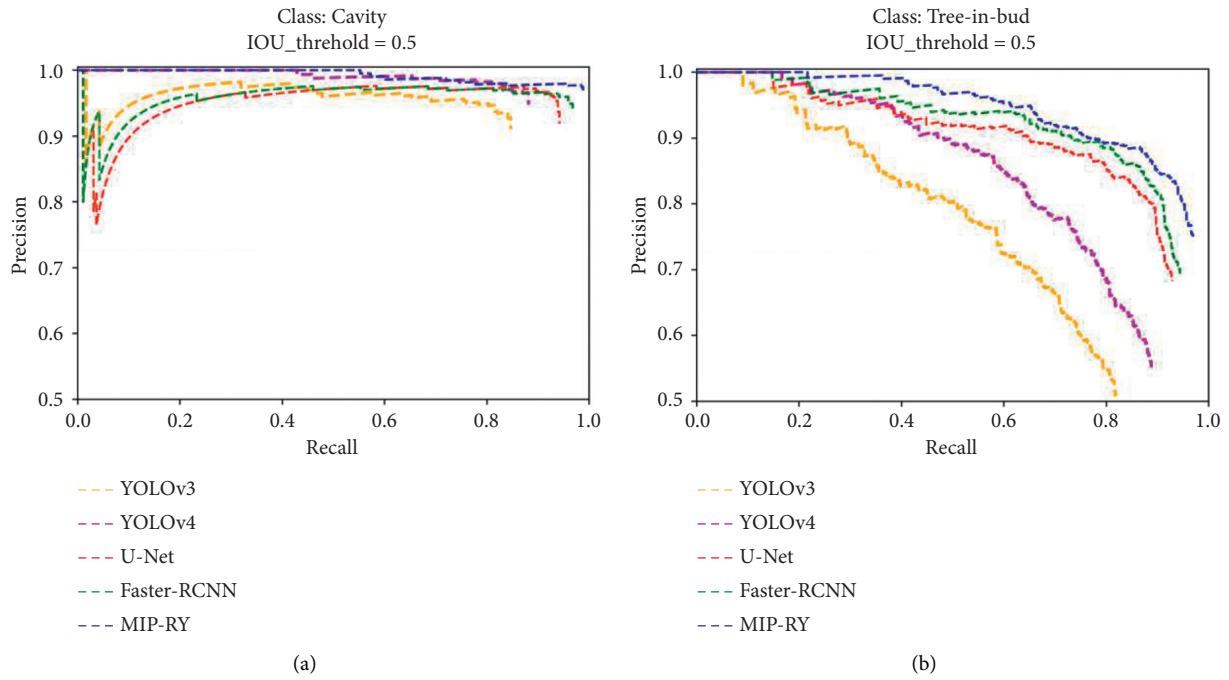


FIGURE 8: P-R curves of different detection models. (a) P-R curves of Cavity; (b) P-R curves of Tree-in-bud pattern.

TABLE 3: Evaluation results of each detection model on test set 2.

Model	Precision (%)	Recall (%)	mAP (%)	Miss detection rate (%)	Parameters (M)	Detection time (s)
U-net [6]	92.20	81.87	88.79	12	96.32	10.36
YOLOv3 [8]	89.14	76.08	74.92	22	61.53	9.63
Faster-RCNN [10]	94.96	83.66	92.40	8	136.65	11.62
YOLOv4 [12]	91.37	81.03	87.21	14	63.94	8.81
MIP-MY	96.59	85.50	95.32	6	33.91	5.72

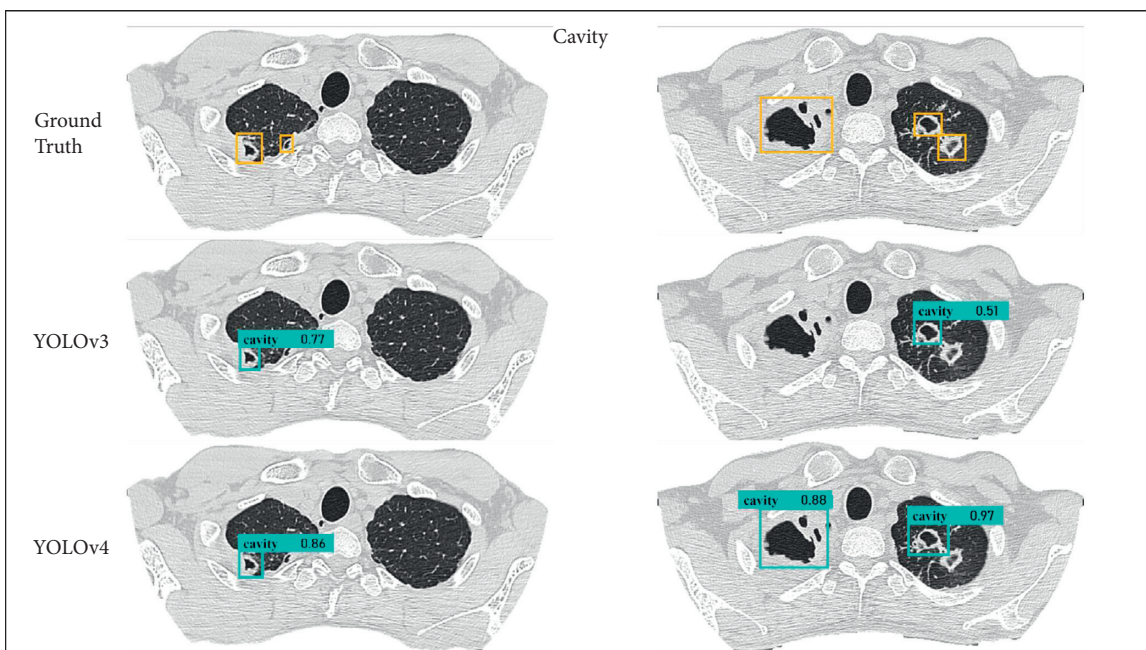


FIGURE 9: Detection of cavity detection results by YOLOv3 and YOLOv4.

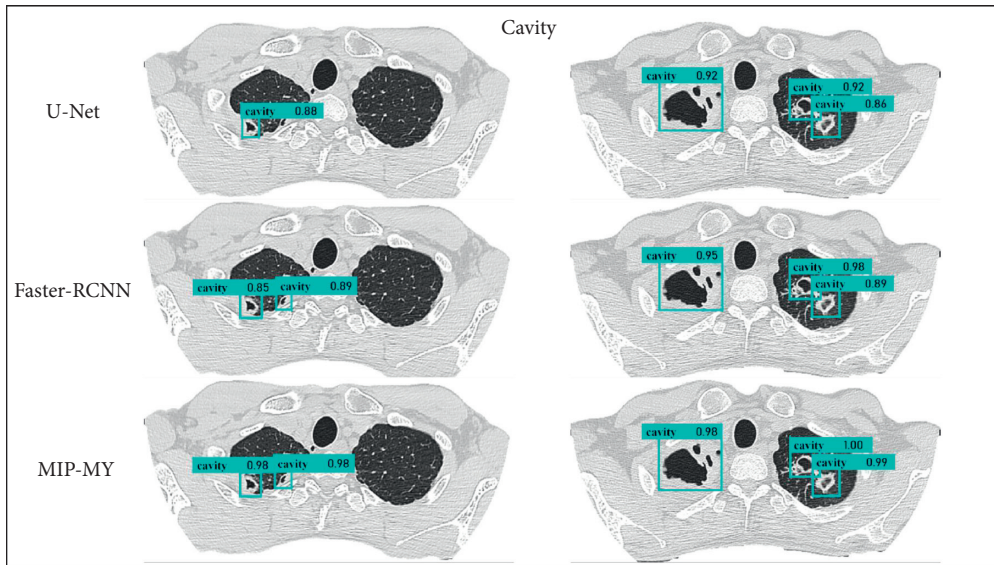


FIGURE 10: Detection of cavity by U-Net, Faster-RCNN and MIP-MY.

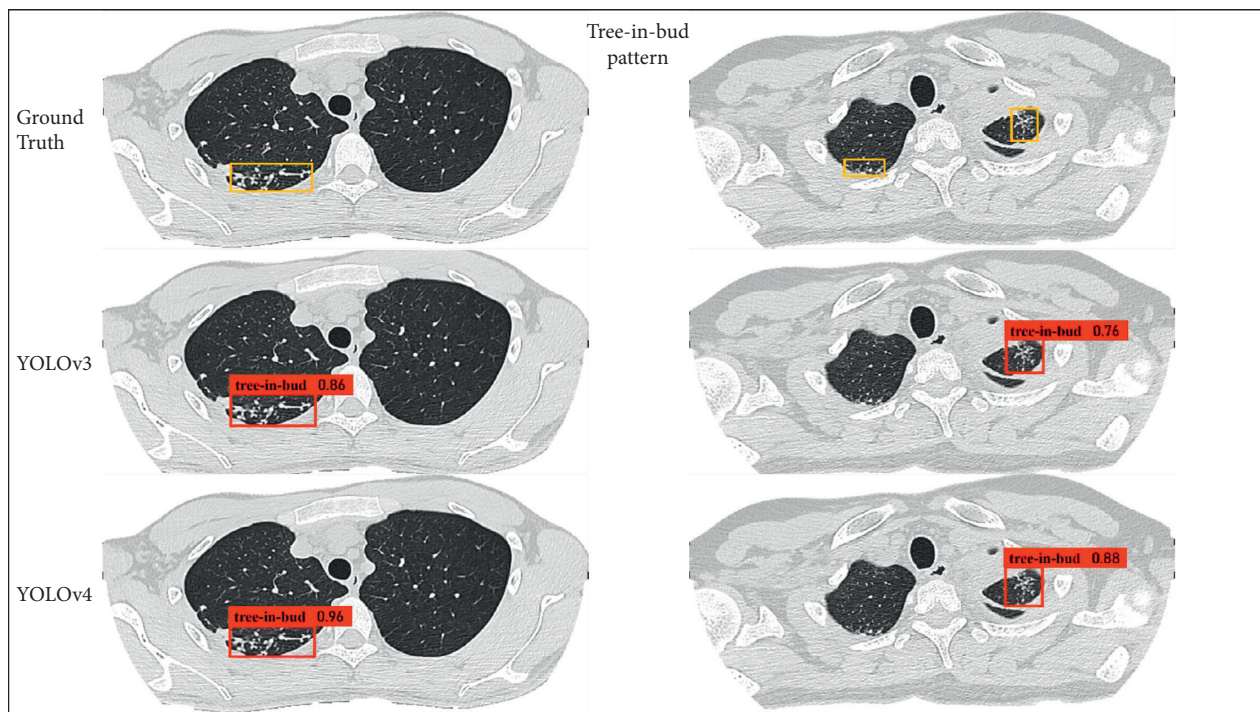


FIGURE 11: Detection of the tree-in-bud pattern by YOLOv3 and YOLOv4.

models is shown in Figure 8. As can be seen from the graph, compared to other mainstream detection models, the scope surrounded by the P-R curve of MIP-MY is the largest. Especially in Figure 8(b), the optimization effect of MIP-MY on the P-R curve for detecting the tree-in-bud pattern is the most obvious, indicating that this model is easier to capture the clustered tree-in-bud pattern.

Table 3 shows the evaluation results of different pulmonary tuberculosis detection models on test set 2. Through quantitative analysis, the performance of the single-stage detection model of YOLOv3 in reference [8]

and YOLOv4 in reference [12] is not up to the mark. The reason is that there are many small cavities and clusters of tree-in-bud pattern in test set 2, which can easily be confused with normal bronchi in the lung region, resulting in a high rate of missed detection in both models. In contrast, the U-Net in reference [6] and the two-stage detection model Faster-RCNN in reference [10] have adequate mAP and low miss detection rate, but these two models have a large number of parameters, so they may not be suitable for equipment with general computing power, and the detection time is slightly



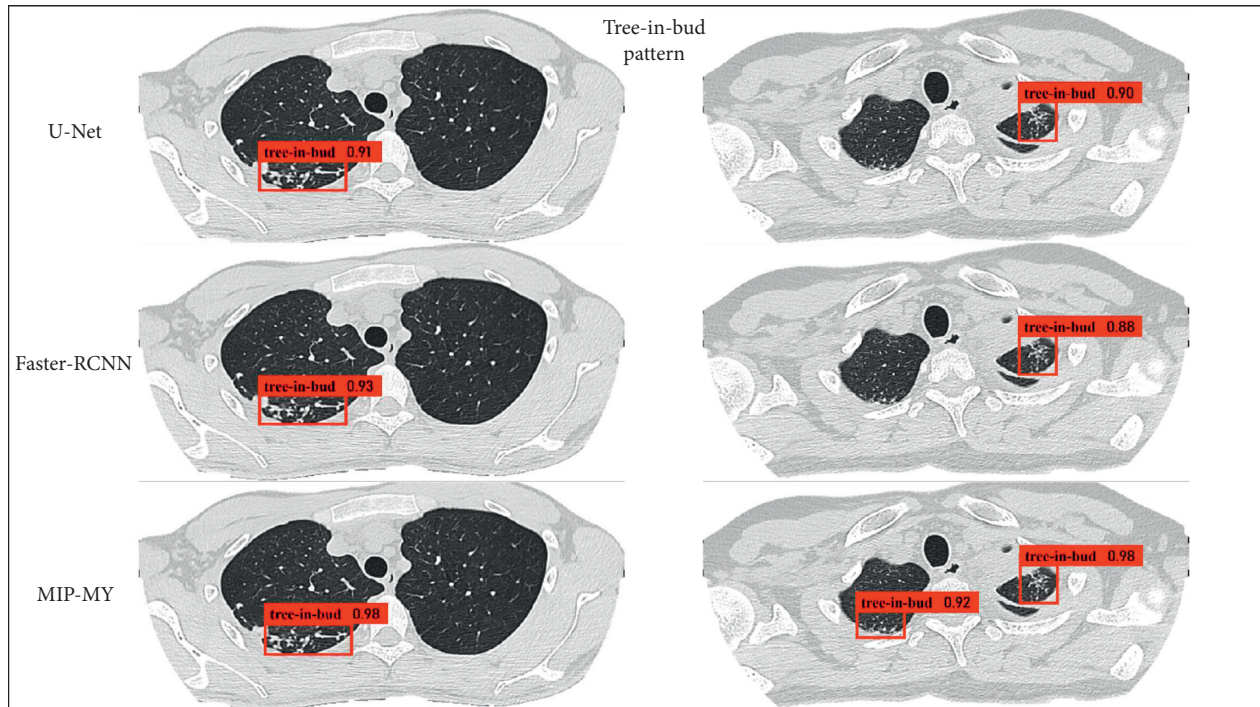


FIGURE 12: Detection of tree-in-bud pattern by U-Net, Faster-RCNN, and MIP-MY.

longer, so it is difficult to achieve the requirements of high efficiency.

The MIP-MY model proposed in this paper has a remarkable performance in test set 2. Compared with the YOLOv4 model, the precision is increased by 5.22%, recall is increased by 4.47%, mAP is increased by 8.11%, the number of parameters is reduced by about 47%, and the error detection rate is reduced by 8%. Furthermore, MIP-MY also has a higher detection efficiency, the detection time of a single pulmonary *tuberculosis* image is shortened by 5.9 s compared to the Faster-RCNN.

From the numerical analysis of the evaluation results, MIP-MY can meet the target of high detection accuracy, and its model parameters take up less space, further reinforces the degree of lightweight.

In this paper, the pulmonary *tuberculosis* detection effect of the proposed MIP-MY model MIP-MY proposed needs to be shown more intuitively. The detection results of cavity in each model are shown in Figures 9 and 10, and the detection results of tree-in-bud pattern are shown in Figures 11 and 12. Ground Truth is the real label prescribed by the imaging doctors, the pulmonary *tuberculosis* cavity is marked as “cavity” in the bounding box, the tree-in-bud pattern is marked as “tree-in-bud” in the bounding box, and the corresponding confidence score is attached to the bounding boxes, which reflects the credibility of these lesions identity.

Through the analysis of the detection results of YOLOv3 and YOLOv4, it is found that the confidence score of YOLOv3 detection cavity and tree-in-bud pattern is low, which may be due to the underutilization of multiscale feature information by YOLOv3. Although YOLOv4 has a high confidence score, it fails to detect small lesions because the down sampling scale of the backbone extractor is too

large and the spatial and pixel information of some small lesions is ignored in the layer-by-layer feature extraction. However, if the down sampling scale is reduced, the detection accuracy of other targets cannot be guaranteed. In Figure 10, it shows that U-Net makes miss detection in a tiny cavity, compared with the ability of Faster-RCNN to accurately capture cavities of various sizes. But in the example in Figure 12, U-Net and Faster-RCNN also omitted the detection in some inconspicuous tree-in-bud pattern.

MRFB module can effectively solve these problems of missed detection about lesions. MIP-MY was observed to capture small cavities and clusters of tree-in-bud patterns that are difficult to identify, indicating that the multiple-receptive field information collection ability of the MRFB module produces a marked effect and avoids the risk of small targets being missed. At the same time, the pulmonary *tuberculosis* lesion detected by MIP-MY has a high-level confidence score, which provides a reliable digital explanation for the automatic diagnosis ability of the detection model.

## 5. Conclusions

In this paper, a lightweight detection model MIP-MY with multiple receptive fields is developed for the automatic detection of pulmonary *tuberculosis*. Through ablation experiments, the contribution of MIP module in model lightweight is verified, and the superior performance of MRFB module and integrated loss function in improving model detection accuracy is also verified. Comparative experiments with other references show that MIP-MY has a lower number of parameters, better detection efficiency and accuracy, and its multiple receptive field characteristics

strengthen the attention to small lesions and greatly reduce the possibility of misdetection. To sum up, the improved model MIP-MY achieves higher detection accuracy, realizes the automatic detection of pulmonary tuberculosis cavity and tree-in-bud pattern with lower calculation cost, and has excellent imaging diagnostic potential of pulmonary tuberculosis. In subsequent studies, more effective data enhancement techniques can be used to enrich the diversity of pulmonary tuberculosis CT data, such as generating game networks (GAN) to generate more samples. In addition, the generality of the model will be studied to make MIP-MY model suitable for the detection of common lung diseases, including lung cancer and pneumonia.

## Data Availability

The dataset used to support the findings of this study has not been made available because the dataset is provided by the partner hospital and includes patient information.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References

- [1] A. Fearn, D. J. Greenwood, A. Rodgers, H. Jiang, and M. Gutierrez, "Correlative light electron ion microscopy reveals in vivo localisation of bedaquiline in lungs infected with Mycobacterium tuberculosis-infected lungs," *PLoS Biology*, vol. 18, no. 12, 2020.
- [2] W. H. Organization, "Global tuberculosis report 2015," *Australasian Medical Journal*, 2016.
- [3] K. I. Yousif, M. E. Maki, R. Babikir, and H. Abuaisha, "Effect of an educational intervention on awareness of various aspects of pulmonary tuberculosis in patients with the disease," *Eastern Mediterranean Health Journal*, vol. 27, 2020.
- [4] A. S. Bhalla, A. Goyal, R. Guleria, and A. Gupta, "Chest tuberculosis: radiological review and imaging recommendations," *Indian Journal of Radiology and Imaging*, vol. 25, no. 3, 2015.
- [5] X. Gao and Y. Qian, "Prediction of multi-drug resistant TB from CT pulmonary Images based on deep learning techniques," *Molecular Pharmaceutics*, vol. 15, 2017.
- [6] L. Ma, Y. Wang, L. Guo, F. Lure, P. Wang, and L. Qian, "Developing and verifying automatic detection of active pulmonary tuberculosis from multi-slice spiral CT images based on deep learning," *Journal of X-Ray Science and Technology*, vol. 28, no. 5, 2020.
- [7] O. Ronneberger, P. Fischer, and T. Brox, *U-net: Convolutional Networks for Biomedical Image Segmentation*, Springer, Berlin, Germany, 2015.
- [8] C. Liu, S.-C. Hu, C. Wang, K. Lafata, and F.-F. Yin, "Automatic detection of pulmonary nodules on CT images with YOLOv3: development and evaluation using simulated and patient data," *Quantitative Imaging in Medicine and Surgery*, vol. 10, no. 10, pp. 1917–1929, 2020.
- [9] J. Redmon and A. Farhadi, "YOLOv3: an incremental improvement," 2018, <https://arxiv.org/abs/1804.02767>.
- [10] A. Yang, X. Jin, and L. Li, "CT images recognition of pulmonary tuberculosis based on improved faster rcnn and u-net," in *Proceedings of the 2019 10th International Conference on Information Technology in Medicine and Education (ITME)*, Qingdao, China, August 2019.
- [11] Z. X. Ye and H. Y. Zhang, "Lightweight improvement of the YOLOv4 mask detection algorithm," *Computer Engineering and Applications*, vol. 57, no. 17, pp. 157–168, 2021.
- [12] A. Bochkovskiy, C. Y. Wang, and H. Liao, "YOLOv4: optimal speed and accuracy of object detection," 2020, <https://arxiv.org/abs/2004.10934>.
- [13] A. Howard, M. Sandler, B. Chen et al., "Searching for mobilenetV3," in *Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, IEEE, Seoul, Korea, November 2020.
- [14] Y. Shang and G. Wu, "Vehicle identification algorithm based on UAV lightweight YOLOV5 snow background," *International Core Journal of Engineering*, vol. 7, 2021.
- [15] X. Zhu, S. Lyu, X. Wang, and Q. Zhao, "TPH-YOLOv5: improved yolov5 based on transformer prediction head for object detection on drone-captured scenarios," 2021, <https://arxiv.org/abs/2108.11539>.
- [16] K. Dheda, L. Lenders, S. Srivastava et al., "Spatial network mapping of multidrug resistant lung tuberculosis cavities using RNA sequence," *American Journal of Critical Care and critical care medicine*, vol. 200, no. 3, 2019.
- [17] A. Sharma, B. Narayanan, N. Pal, and S. Taye, "An unusual cause of 'tree-in-bud' appearance in CT-chest during COVID-19 pan-demic," *Monaldi Archives for Chest Disease = Archivio Monaldi per le Malattie del torace*, vol. 90, no. 3, 2020.
- [18] A. A. Hidayat, T. W. Cenggoro, and B. Pardamean, "Convolutional neural networks for scops owl sound classification," *Procedia Computer Science*, vol. 179, no. 4, pp. 81–87, 2021.
- [19] C. H. Tu, J. H. Lee, Y. M. Chan, and C. Chen, "Pruning depthwise separable convolutions for mobilenet compression," in *Proceedings of the 2020 International Joint Conference on Neural Networks (IJCNN)*, Budapest, Hungary, July 2020.
- [20] H. Wang, F. Zhang, and L. Wang, "Fruit classification model based on improved darknet53 convolutional neural network," in *Proceedings of the 2020 International Conference on Intelligent Transportation*, Vientiane, Laos, January 2020.
- [21] P. Purkait, C. Zhao, and C. Zach, "SPP-net: deep absolute pose regression with synthetic views," in *Proceedings of the British Machine Vision Conference (BMVC 2018)*, Newcastle, UK, April 2017.
- [22] S. Liu, L. Qi, H. Qin, and J. Jia, "Path aggregation network for instance segmentation," in *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Salt Lake City, UT, USA, June 2018.
- [23] H. Zhao, J. Shi, X. Qi, and X. Wang, "Pyramid scene parsing network," in *Proceedings of the IEEE Computer Society*, Honolulu, HI, USA, July 2016.
- [24] S. Liu, D. Huang, and Y. Wang, "Receptive field block net for accurate and fast object detection," in *Proceedings of the European Conference on Computer Vision*, Munich, Germany, September 2018.
- [25] A. S. Yan, Y. L. Xin, and G. Liang, "Improved non-maximum suppression for detecting overlapping objects," in *Proceedings of the International Conference on Machine Vision*, Yangon, Myanmar, November 2020.
- [26] M. Christine, G. Daniel, H. Tim, S. Denise, S. Adrian, and S. Tobias, "Anaphora resolution: analyzing the impact on mean average precision and detecting limitations of automated approaches," *International Journal of Information Retrieval Research*, vol. 8, no. 3, 2018.

## Research Article

# Simulation Study on Complex Systems of Forest Biomass Power Generation Supply Chain in China

Ning Ma , Zhao Huang , and Yue Qi 

*School of Economics & Management, Beijing Forestry University, Beijing 100083, China*

Correspondence should be addressed to Ning Ma; [maning@bjfu.edu.cn](mailto:maning@bjfu.edu.cn)

Received 26 October 2021; Accepted 21 February 2022; Published 29 March 2022

Academic Editor: Daqing Gong

Copyright © 2022 Ning Ma et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Based on the perspective of complex system management of forest biomass power generation (FBPG) supply chain, this paper explores a new way to reduce the operation cost of forest biomass raw materials supply chain and improve the operation efficiency of forest biomass supply chain in China. Firstly, the supply chain model of FBPG is summarized through the survey data, and the existing problems are found. Then, the behavior strategies among farmers, middlemen, and enterprises are analyzed, and the data that cannot be obtained in the model are estimated by using statistical methods. Finally, the complex supply chain system is simulated by computer. Through setting different forestry policies, this paper compares and analyzes the profit of the supply chain and puts forward policy suggestions to promote the development of the FBPG industry in China. In this paper, the actual situation, principles of statistics, and system simulation are effectively combined to make the behavior rules of various stakeholders fit the reality, and the research conclusions are more practical.

## 1. Introduction

With the development of society, the importance of the ecological environment has become increasingly prominent. As a renewable and clean energy industry, biomass energy industry is of great significance to improve environmental quality and develop circular economy. Forest biomass power generation (FBPG) industry is an important part of biomass energy industry. It can generate economic benefits by making rational use of forestry resources and bring substantial ecological benefits by reducing pollutant emissions to protect the natural environment. Especially for the northwest area with poor natural conditions in China, the energy development of tending residues such as pruning shrubs can improve the income level of local farmers and the local natural environment.

Although the utilization of forest biomass energy is essential, China's FBPG industry is still in the initial stage of development, and there are many problems in actual development. For example, farmers lack the will to collect,

so it is difficult to collect forest biomass raw materials on a large scale. In addition, the cost of raw materials remains high, and the degree of marketization is low. Therefore, from the perspective of the complex system of forest biomass raw materials supply chain, it is of great significance to mobilize the enthusiasm of farmers to collect and reduce the cost of raw materials by optimizing the policy combination so as to maximize the overall profit of the supply chain and ensure the sustainable and stable operation of biomass power generation enterprises.

This paper combines the supply chain of FBPG with computer simulation technology. Firstly, the paper models the complex supply chain system composed of farmers, middlemen, and biomass power generation enterprises and constructs a multiagent raw materials supply chain simulation model. Then, the paper puts forward the optimal combination scheme and policy suggestions for the development of the FBPG industry through simulating the simulation results under different parameters, which has important practical significance for expanding the scale of the FBPG industry in China.



## 2. Literature Review

Forest biomass refers to the organic matter formed by photosynthesis of woody plants such as forest trees. Forest biomass energy is the energy stored in forest biomass after conversion; that is, the energy can be used for power generation or heating through direct combustion or modern conversion technologies [1, 2]. Forest biomass resources can be divided into three categories: forest tending and pruning residues, “three forestry residues” (cutting residues, material residues, and processing residues), and energy forest mining residues. The overall availability of forest biomass energy resources in China is increasing year by year, and the supply capacity is significant. In particular, Inner Mongolia has the most extensive available forest biomass energy resources, accounting for 9% of China [3]. The forest biomass energy industry in the western region has great development potential.

The development of the biomass energy industry still faces problems such as difficult raw materials collection, low equipment utilization, and high operation and maintenance costs [4, 5]. By 2016, the global consumption of biomass power generation accounted for about 40% of the total consumption of biomass energy [6]. Energy has not been fully utilized, which restricts the sustainable development of biomass energy industry to a great extent. Insufficient purchase of raw materials is one of the important factors limiting the development of forest biomass energy industry [7]. At present, China’s forest biomass energy mainly depends on government policy support [8]. It is urgent to explore the enthusiasm of various agents from the internal system and put forward corresponding incentive measures.

The simulation modeling method based on multiagent started in the late 1990s. This method adopts bottom-up modeling thought and emphasizes the authenticity of microindividuals during modeling. The model establishes the relationship between agents through the message mechanism and then forms the macroframework of the simulation system. Many scholars have applied multiagent simulation technology to the supply chain of forestry, agricultural products, and other fields. For example, Hanafizadeh and Sherkat [9] built a supply chain distribution model based on multiagent system, improved the convergence of agent adaptability by introducing a genetic algorithm, and then studied the related problems of distribution in supply chain management. Wang et al. [10] used NetLogo software to build a simulation model of agricultural product supply chain and explored the relationship between supervision, farmers’ cooperation, and agricultural product quality. Zhang and Lin [11] established a simulation model based on multiagent to study the impact of the subsidy policy issued by the Chinese government on the profit distribution of supply chain members and carried out experiments under different subsidy levels. Terrada and Ouajji [12] applied multiagent simulation technology to the complex system of supply chain management to provide solutions to the decision-making problems among supply chain members. Awaga et al. [13] built a simulation analysis model of enterprise green product production behavior with multiagent

participation and simulated the impact of different government guidance and supervision strategies on the strategic choices of two parties (enterprises and government). In addition, few scholars have studied biomass power generation industry. Shastri et al. [14] simulated biomass supply system in Illinois based on multiagent simulation technology, and the simulation revealed the impact of miscanthus energy crop on raw materials supply contract. Li [15] established an industrial system model including farmer, enterprise, market, and government and put forward suggestions on the construction scale of China’s straw power generation and the centralized utilization of straw. Luo [16] constructed four biomass raw materials supply game models based on game theory and multiagent simulation technology. Then he analyzed the effect, applicability, and corresponding implementation strategies of biomass power generation raw materials supply mode.

Because of the different characteristics of supply chains, there are also differences in agent behavior rules. The multiagent modeling of the FBPG supply chain is more considered from the perspective of raw materials supply. According to the research results, farmers, middlemen, and enterprises can be abstracted as agents with independent decision-making, learning, and memory ability. There are obvious differences in the decision-making rules of agents, the interaction rules between agents and between agents, and the external environment. The use of multiagent simulation technology shows obvious advantages. The behavior rules of simulation agents are designed according to the actual behavior of individuals. Through the interaction of microagents, various emergent behaviors and self-organization phenomena in the system can be effectively explored [17].

In conclusion, the forest biomass energy industry has extensive development space. The existing problems are mainly backward technology, lagging policy, insufficient supply of raw materials, and high production cost. Therefore, it is necessary to study the supply chain of forest biomass energy raw materials from the perspective of microagents. Policy factor greatly impacts biomass power generation industry, which can promote industrial development through economic policy adjustment. In addition, although the multiagent simulation method is not widely used in the forest biomass energy supply chain, there have been many studies in the field of forestry and agriculture. Therefore, it is reasonable to summarize the industrial policy suggestions through a series of simulation experiments.

## 3. Forest Biomass Power Generation Supply Chain

*3.1. Characteristics of Supply Chain.* The characteristics of FBPG supply chain are as follows:

- (1) As a biomass enterprise, the power generation product cannot be saved.
- (2) Raw materials have a wide production area and large output, but the collection has periodicity and is greatly affected by factors such as environment and season.

- (3) There are many alternative uses of raw materials, resulting in large supply fluctuation.
- (4) Raw materials are susceptible to moisture and decay, with low weight and value density. The cost of transportation and storage is high.
- (5) The raw materials supply of enterprise is generally subject to many farmers and middlemen. The enterprise does not occupy a dominant position.

It can be seen that FBPG supply chain is essentially different from the conventional production supply chain. The latter usually only focuses on the uncertainty of the consumption market, while the former pays more attention to the uncertainty of upstream raw materials supply.

In addition, compared with biomass raw materials such as straw, the geographical distribution of forest biomass raw materials is more special, the collection quantity is more uncertain, and the transportation cost is more expensive. Therefore, the upstream of FBPG supply chain is relatively complex, and the supply is unstable. By signing contracts and expanding the scope of purchase, the cooperative relationship between various agents can be stabilized. When establishing the simulation model, seasonal and other random factors will be set to reflect supply uncertainty.

**3.2. Composition of Supply Chain.** The first-hand data used in this paper was obtained by issuing questionnaires in Wushen Banner and Hangjin Banner of Inner Mongolia Autonomous Region, and Zhongwei City of Ningxia Hui Autonomous Region, as well as in-depth interviews with enterprises' leaders. The survey covers the basic information and raw materials purchase information of enterprises. In this survey, 5 enterprises were visited, and 127 questionnaires of farmers related to biomass energy were completed. Fifty-four farmers have sold raw materials in two ways (Table 1).

In the upstream supply chain of FBPG, there are mainly two kinds of raw materials transportation modes. One is that forest farmers directly sell to FBPG enterprises; the other is that forest farmers sell to middlemen and then sell to enterprises by middlemen. The downstream supply chain is relatively simple. The only customer is the State Grid Corporation of China (SGCC). The whole FBPG supply chain system can be shown in Figure 1. Next, the supply chain is divided into cost supply chain and sale supply chain to analyze the upstream and downstream customers outside the enterprise.

**3.2.1. The Upstream Supply Chain.** The upstream supply chain of FBPG enterprises in China is mainly two-level or three-level supply chain: (1) two-level supply chain: farmers/upstream companies/enterprises (independently cultivated raw materials)-power generation enterprises; (2) three-level supply chain: farmers-middlemen-power generation enterprises. Specifically, power generation enterprises mainly obtain raw materials through four modes: first, power generation enterprises independently cultivate economic forests as raw materials for power generation; second,

enterprises cooperate with upstream companies to obtain raw materials, which are generally molding fuel enterprises; third, enterprises purchase crops or forest residues from farmers; fourth, enterprises purchase raw materials from middlemen. These middlemen are mainly individual farmers. Due to the market economy, they have spontaneously become a hub connecting upstream and downstream raw materials.

The choice of raw materials supply mode depends on many factors such as business plan, capital status, geographical location, and other factors. Enterprises often choose mixed modes. Mode 1 belongs to internal production of enterprises and has a low degree of application. Mode 2 is interest cooperation among enterprises. Mode 3 and mode 4 are transactions between individual farmers and enterprises. Therefore, this paper mainly studies mode 3 and mode 4.

**3.2.2. The Downstream Supply Chain.** The downstream supply chain of FBPG is a two-level supply chain: power generation enterprises-SGCC. In addition to a few enterprises using the advantage of biomass power generation to increase sideline income, SGCC is the only downstream customer in the sale chain of the whole industry. Therefore, the formulation of policies on FBPG largely determines the survival and development of power generation enterprises.

At present, the feed-in tariff standard of biomass energy in China is 0.75 yuan/kWh. The feed-in tariff is the price of power commodity delivered to SGCC by power production enterprises. As China has formulated the priority scheduling policy, the sale of biomass power generation will be guaranteed. Moreover, according to the Renewable Energy Law, biomass power is preferentially connected to the grid and does not participate in peak regulation. Therefore, these policies reduce the impact of changes in electricity consumption of end customers on biomass power generation industry and ensure the stability of sale of biomass power generation enterprises.

## 4. Construction of the Simulation Model

This part establishes the three-level supply chain simulation model consisting of farmers, middlemen, and biomass power generation enterprises and uses NetLogo software for simulation. The role of the government appears in the form of subsidy parameters in the simulation system. The behaviors of farmers and middlemen adopt the "economic man" hypothesis. Middlemen will adjust their purchase scope according to the price of enterprise to maximize their profit. The simulation system takes enterprise as the core node of the whole supply chain to simulate the optimal purchase price and government subsidy when the enterprise reaches production capacity. The parameters and some processes of the model are mainly based on the relevant data of the Maowusu biomass thermal power plant in Erdos, Inner Mongolia. The enterprise takes the branches of sandy shrubs (mainly *Salix psammophila*) pruned from the roots as raw materials.

TABLE 1: Ways for farmers to sell raw materials.

Selling mode of farmers	Farmers sell to enterprises	Farmers sell to middlemen
Farmers take the initiative to transport	19	24
Farmers waiting to be acquired	4	7
Total	23	31

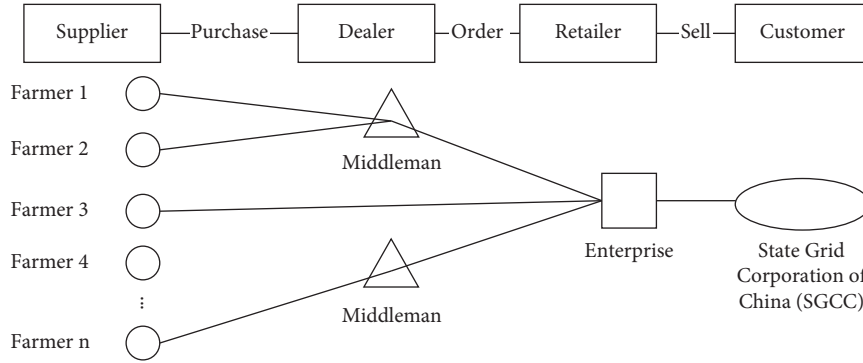


FIGURE 1: The supply chain system of FBPG.

*4.1. Simulation Assumptions and Process.* Under the premise of ensuring the authenticity of the system as much as possible, the simulation model is simplified, and the following assumptions are made according to the actual situation:

- (1) There is only one biomass power generation enterprise in a particular region, and there is no competition for raw materials among multiple biomass power generation enterprises.
- (2) During the simulation period, the area of woodland and the number of labor forces in the region do not change, and the positions of biomass power generation enterprise, middlemen, and farmers do not change.
- (3) The enterprise has no difficulty in selling products; that is, all electricity can be sold, and the enterprise is only responsible for purchase, not transportation.
- (4) According to the current policy, the government implements the tax policy of “immediate levying and refunding,” so the impact of bank tax is not considered in the model.
- (5) Farmers know the purchase price of middlemen and enterprise last time before pruning. The behaviors of farmers adopt the “economic man” hypothesis.
- (6) The distribution of forest biomass resources is extensive, uniform, and unitary. The woodland area is infinite. The species are single, and all of them are *Salix psammophila*. They are evenly distributed in the region.
- (7) Middlemen take the initiative to purchase raw materials from farmers’ homes.

In the simulation processes of FBPG supply chain, firstly, initialize the system state to generate the farmer agent, the

middleman agent, and the biomass power generation enterprise agent. Secondly, the enterprise sets the price for purchasing raw materials, and middlemen set a reasonable purchase range according to the current price. Thirdly, farmers decide whether to collect and sell forest biomass raw materials according to the purchase price of enterprise and middlemen. If the farmers sell to middlemen, the middlemen are responsible for transportation and primary processing and then sell to the enterprise; if farmers sell directly to the enterprise, farmers need to pay the cost of transportation. After purchasing raw materials, the enterprise processes and generates electricity and then sells electricity to SGCC to obtain government subsidy. Finally, calculate the profit of the biomass power generation enterprise. Judge whether the quantity of purchase can meet the demand of enterprise. If the enterprise reaches the production capacity, the program will stop; if the production capacity is not reached, the enterprise will increase the purchase price until the program stops when the production capacity is reached. At the same time, calculate the profit of farmers and middlemen in each cycle.

The simulation system finally solves the following three problems: (1) How much purchase price can make the biomass power generation enterprise reach production capacity and obtain the maximum ecological benefits. (2) How much government subsidy can ensure that the enterprise does not lose money at the purchase price of reaching production capacity. (3) What policy can be implemented to improve the overall efficiency of the supply chain.

#### 4.2. Construction of the Simulation Model

*4.2.1. Property and Parameter Setting of the Farmer Agent.* According to the survey data of Inner Mongolia, about 140 farmers directly transport raw materials to the enterprise

every year, involving 100–300 kilometers. About 20 middlemen transport raw materials to the enterprise every year, also involving 100–300 kilometers. Each middleman purchases raw materials from about 20 farmers. Therefore, there are about 540 farmers involved in selling raw materials. At the same time, in the farmers' questionnaires, the number of farmers who have sold raw materials in recent three years accounts for 48.1% of the respondents. It is roughly estimated that the number of forest farmers within 300 kilometers of the enterprise is about 1123. Moreover, because farmers prune every three years, it is estimated that the number of farmers is about 3369. Each  $1/15 \text{ hm}^2$  *Salix psammophila* can produce about 1t branches [18]. However, pruning *Salix psammophila* is greatly affected by external factors such as season and family labor force. The actual unit collection is often less than 1t. Therefore, a random variable  $\alpha$  (value between 0.5~1) is designed to more truly reflect the pruning situation of farmers. The pruning cost is mainly labor cost, and the pruning cost per ton of raw materials is about 70~200 yuan.

According to the survey, the woodland area of farmers ranges from 5 to 5400 mu. Through the K-S test, the woodland area data is not a normal distribution, and there are obvious abnormal values. Clean the woodland area data, sort the data from small to big, and take 25% as the upper quartile FL and 75% as the lower quartile Fu. The quartile distance IQR is Fu-FL. Then the upper bound of the data is  $Fu + 1.5 \times \text{IQR}$ , the lower bound is  $FL - 1.5 \times \text{IQR}$ , and the data ranges from 5 to 1585 mu after cleaning.

Grouping the data with 200 as group interval, it can be found that the data distribution is close to the exponential distribution. After fitting the exponential distribution, the fitting function and goodness of fit are obtained, and the fitting effect is good, as shown in Figure 2.

Based on the previous analysis and survey data, the main parameters of the farmer are sorted into Table 2.

**4.2.2. Property and Parameter Setting of the Middleman Agent.** Middlemen purchase branches from farmers but are not responsible for the harvesting and collecting branches. The purchase price is 180 yuan/ton. The cost of raw materials primary processing (including chipping and bundling) ranges from 20 yuan/ton to 35 yuan/ton. The main parameters of the middleman are sorted into Table 3.

#### 4.2.3. Property and Parameter Setting of the Enterprise Agent

- (1) Calculation of raw materials demand of the enterprise: taking Maowusu power plant as an example, the unit installed capacity is  $2 \times 12 \text{ MW}$ , and the power generation per hour is  $2 \times 1.2 \text{ kW}$ . As pruning and collecting of *Salix psammophila* occur on the sand, the raw materials after chipping contain a large amount of sand. It is necessary to shut down for 20–25 days every year to clean the tube wall; otherwise, the pipeline will be blocked and the service life of the unit will be reduced. Therefore, if the generator unit reaches production capacity,

assuming that it needs to be shut down for 25 days, the annual power generation is 195.84 million ( $2 \times 1.2 \times 24 \times 340$ ) kWh. According to the survey data, 1.25 Kg *Salix psammophila* can be converted into 1 kWh, so at least 156672 tons ( $19584 \times 10^4 / 1.25 / 10^3$ ) *Salix psammophila* can meet the needs of the enterprise. In fact, considering that the enterprise has other raw materials besides *Salix psammophila*, the demand is about 100000 tons. (2) Cost and profit: at present, the feed-in tariff of SGCC is 0.75 yuan/kWh, of which 0.277 is the benchmark price. Without considering the bank cost, the power generation cost per kWh is 0.292 yuan. The main parameters of the enterprise are sorted into Table 4.

**4.2.4. Behavior Rules of the Farmer Agent.** Farmers aim to maximize their profit and choose a reasonable sale object. Farmers' profit is sale income plus government subsidy and minus all costs. Farmers need to pay the collection cost of collecting biomass raw materials; the income of farmers mainly comes from biomass raw materials in their woodland. If it is sold directly to the enterprise, it also needs to pay the transportation cost. The main influencing factors of biomass raw materials transportation cost include vehicle technical performance, vehicle operating cost, characteristics of crops, collection mode, and radius [19]. Because the actual route is generally curved, the road tortuosity factor  $\lambda$  ( $\lambda = \sqrt{2}$ ) is introduced for correction. At the same time, the round-trip transportation is considered, so the distance should be twice the actual distance. Therefore, some formulas of farmers' behaviors are as follows:

- (1) Quantity of raw materials collected ( $Q$ ) =  $FS \times UAmo \times \alpha$ .
- (2) Pruning cost ( $FC_1$ ) =  $FCost \times Q$ .
- (3) Transportation cost of the farmer  $FC_2 = 2\sqrt{2} \times FE_d \times$  transportation cost per unit distance.
- (4) The transportation cost per unit distance varies with the collection quantity. According to the questionnaires, the type and performance of vehicles mainly used by farmers are shown in Table 5.

It is known that the average oil price in Inner Mongolia is 5 yuan/L. The transportation cost per unit distance of farmer with different collection quantity can be expressed as follows:

When  $0 < Q \leq 0.5$ , the tricycle is used for transportation, and the transportation cost per kilometer is 0.5 yuan.

When  $0.5 < Q \leq 1.5$ , the tractor is used for transportation, and the transportation cost per kilometer is 0.75 yuan.

When  $Q \geq 1.5$ , the four-wheel vehicle is used for transportation, and the transportation cost per kilometer is  $[Q/2.5]$  yuan.

$[Q/2.5]$  means to take the smallest integer larger than itself.

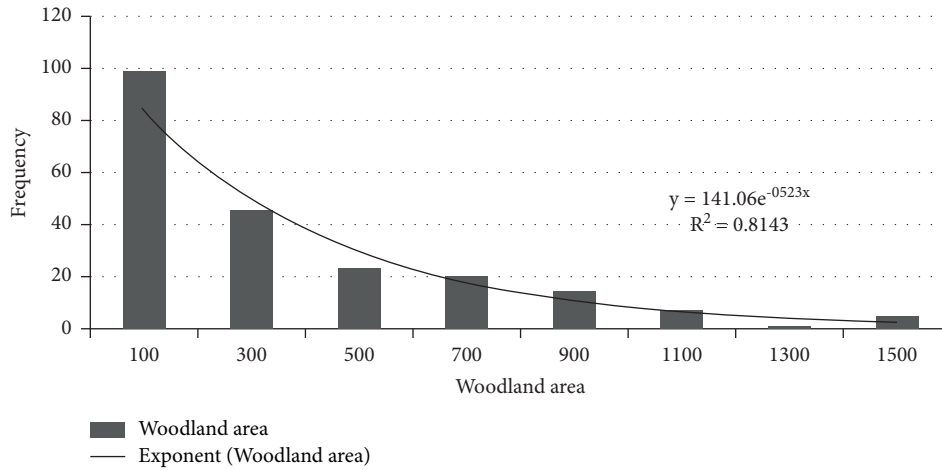


FIGURE 2: Distribution of the woodland area.

TABLE 2: Property and parameter setting of farmers.

Property	Letter representation	Meaning	Assigned value
Woodland area	FS	Woodland area owned by each farmer	$141.06e^{-0.523x}$ mu
Farmer-enterprise distance	FEd	Distance from farmer to biomass enterprise	0~300 km uniform distribution
Interval time of pruning	t	Time since last pruning	0~5 years random distribution
Number of farmers	FNum	Number of farmers in the region	1000~4000 (sliding strip)
Unit amount collected	UAmo	Quantity of raw materials collected per unit woodland	1 t/mu
Actual collection rate	$\alpha$	Random factors such as season	0.5~1 random distribution
Unit pruning cost	FCost	Cost of raw materials collected per unit woodland	70~200 yuan/t (sliding strip)
Unit pruning subsidy	CSub	Government subsidy for unit woodland pruning	Input box input (yuan)

TABLE 3: Property and parameter setting of the middleman.

Property	Letter representation	Meaning	Assigned value
Middleman-enterprise distance	MEd	Distance from middleman to enterprise	100~300 km uniform distribution
Number of middlemen	MNum	Number of middlemen in the region	10~50 (sliding strip)
Middlemen purchase price	MPri	Price of raw materials purchased from farmer by middleman	150~200 yuan (sliding strip)
Unit processing cost	MCost	Raw materials primary processing cost	20~35 yuan/t (sliding strip)

TABLE 4: Property and parameter setting of the enterprise.

Property	Letter representation	Meaning	Assigned value
Enterprise purchase price	EPri	Purchase price of raw materials for enterprise	350 yuan/t (initial value)
Production capacity of the enterprise	EQua	Raw materials required for power generation when reaching production capacity	100000t
Unit power generation cost	ECost	Power generation cost of enterprise	0.292 yuan/kWh
Feed-in tariff	Price	Sale price per kWh	0.55~1.55 yuan/kWh (sliding strip)
Purchase subsidy	GSub	Subsidy price of raw materials per cycle when production capacity is not reached	10 yuan/t
Raw materials conversion rate	$\mu$	How much power is generated per ton of raw materials	800 kWh/t

TABLE 5: Type and performance of farmers' vehicles.

Vehicle type	Tricycle	Tractor	Four-wheel vehicle
Fuel consumption per 100 km (L)	10	15	20
Unit quantity of transportation (t/vehicle)	0.5	1.5	2.5

Therefore, the transportation cost of the farmer ( $FC_2$ ) can be calculated.

- (5) The profit of farmer from selling raw materials to the enterprise is

$$W_1 = EPri \times Q + FS \times CSub - FC_1 - FC_2. \quad (1)$$

- (6) The profit of farmer from selling raw materials to the middleman is

$$W_2 = MPri \times Q + FS \times CSub - FC_1. \quad (2)$$

- (7) The farmer compares  $W_1$  and  $W_2$  to choose to prune or not, and whether to sell to the middleman or sell to enterprise. Calculate the quantity of raw materials actually transported to the enterprise ( $Q_{fe}$ ) and the quantity of raw materials actually transported to the middleman ( $Q_{fm}$ ) in each cycle.

The specific behavior rules of farmer are shown in Figure 3.

It should be noticed that *Salix psammophila* matures in about three years, and the more it is cut, the more lush it is. If the grown branches are not cut off, it will become dead branches in less than seven years. Because the government has public welfare forest subsidy for energy forest, once *Salix psammophila* dies, farmers will not be able to obtain public welfare forest subsidy, which will cause great losses to farmers. Therefore, the system assumes that farmers must choose to prune in the cycle, and the frequency of pruning is set to one time every three years. Moreover, the purchase scope of the middleman may not cover all farmers, so the farmer needs to judge whether he/she is within the purchase scope of the nearest middleman in advance so as to make the next decision of choosing to prune or not and selecting sale object.

**4.2.5. Behavior Rules of Middleman Agent.** Middlemen also aim to maximize their profit and hope to acquire and sell as many raw materials as possible. Middlemen are responsible for the initial processing and transportation of raw materials in the supply chain. The transportation route consists of two parts: one is that the middleman purchases from the farmer and then transports raw materials to the processing point (the location of the middleman), and the other is that the middleman transports raw materials from the processing point to the biomass power generation enterprise. Middleman decides the current purchase scope based on the current purchase price of the enterprise.

According to the questionnaires, there are two main types of vehicles used by middlemen, as shown in Table 6. It is assumed that a large four-wheel vehicle is mainly used when

middleman purchases raw materials from retail farmers, and a truck is mainly used to transport the biomass raw materials after chipping to the enterprise. Because the middleman can purchase from multiple households at the same time, it can be seen that the vehicle is always fully loaded during the purchase of raw materials from retail farmers:

- (1) When middleman purchases raw materials from retail farmers, the unit transportation cost is 0.25 yuan/ton-km (the calculation formula is  $25 \times 5/100/5$ ).

Because the middleman cannot accurately estimate the purchase quantity before determining the purchase scope, the unit transportation cost transported to the enterprise is temporarily calculated as full load so as to calculate the purchase radius  $R$  of the middleman.

- (2) The estimated unit transportation cost of transporting raw materials to the enterprise is 0.067 yuan/ton-km (the calculation formula is  $40 \times 5/100/30$ ).
- (3) The purchase radius  $R$  of the middleman should base on the economy of cost. According to the calculation method of purchase radius in reference [20], the unit cost of the middleman should be less than the purchase price of the enterprise ( $EPri$ ). The road tortuosity factor and round-trip transportation which are the same as farmers are considered. The formula is expressed:  $MPri + MCost + 2\sqrt{2} \times 0.25 \times R + 2\sqrt{2} \times 0.067 \times MEd \leq EPri$

The maximum purchase radius of the middleman is

$$R = \frac{EPri - MPri - Mcost - \sqrt{2} \times 0.133 \times MEd}{\sqrt{2} \times 0.5}. \quad (3)$$

- (1) The purchase quantity of the middleman is the sum of the quantity of raw materials within the purchase scope:

$$Q_m = \sum Q_{fm}. \quad (4)$$

- (2) The primary processing cost of the middleman is  $MC_1 = MCost \times Q_m$ .
- (3) The actual transportation cost of the middleman is

$$MC_2 = 2\sqrt{2} \times \left( 0.25 \times R + \left[ \frac{Q_m}{30} \right] \times 2 \times MEd \right). \quad (5)$$

- (4) The profit of middleman is

$$W_3 = (EPri - MPri) \times Q_m - MC_1 - MC_2. \quad (6)$$

The specific behavior rules of middleman are shown in Figure 4.



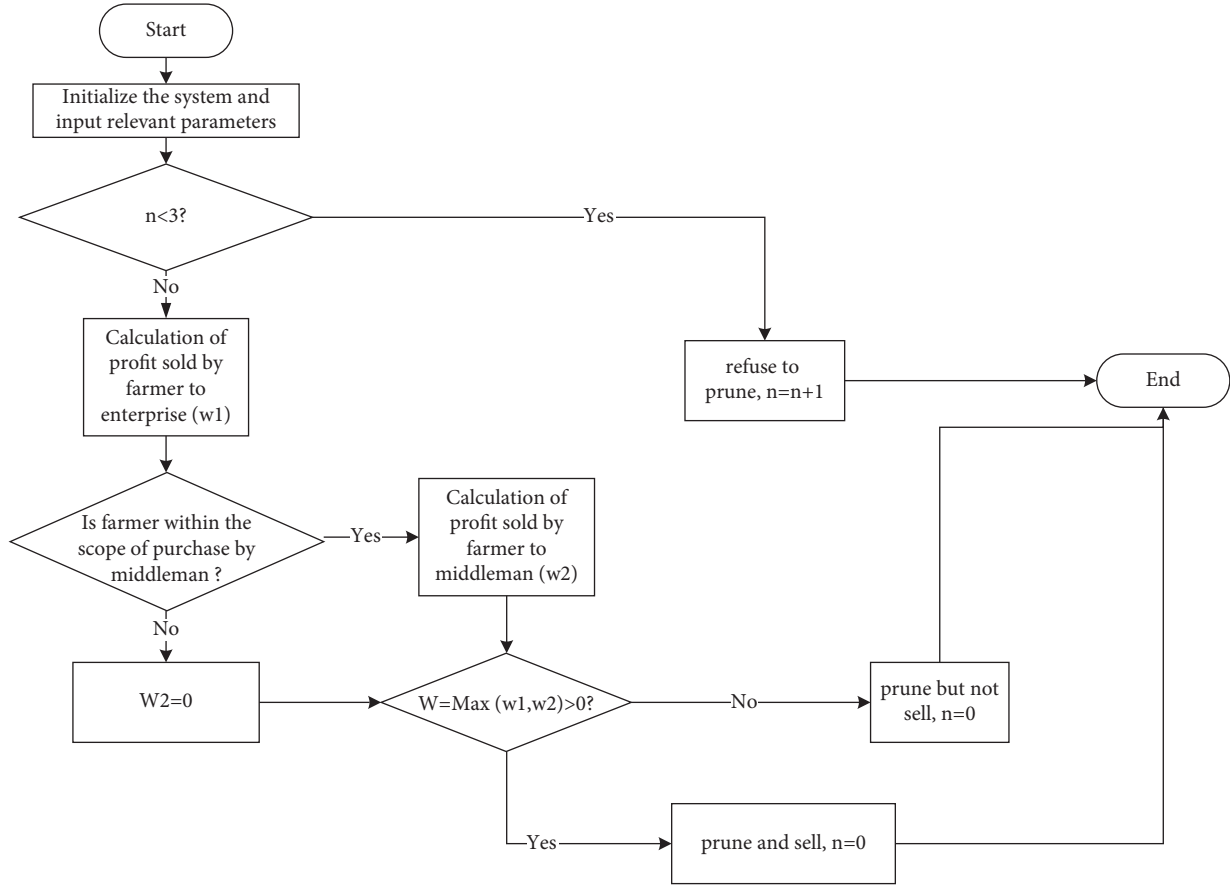


FIGURE 3: Flowchart of farmer's behavior.

TABLE 6: Vehicle type and performance for middlemen.

Vehicle type	Large four-wheel vehicle	Truck
Fuel consumption per 100 km (L)	25	40
Unit quantity of transportation (t/vehicle)	5	30

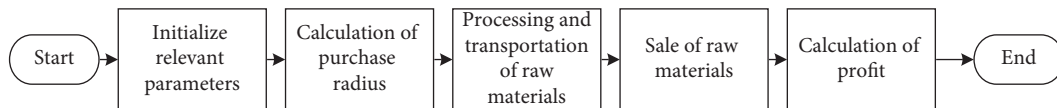


FIGURE 4: Flowchart of middleman's behavior.

**4.2.6. Behavior Rules of the Enterprise Agent.** The enterprise plays a leading role in the whole supply chain. The change of enterprise's purchase price will have an impact on the purchase scope of middlemen and farmers' willingness to sell. The initial price of the enterprise is set according to the current market price. If the production capacity is not reached, the purchase price of the enterprise in the next cycle will increase by 5 yuan/ton. If output exceeds demand, the price in the next cycle will be reduced by 5 yuan/ton to reach an equilibrium state.

Some formulas are calculated as follows:

- (1) Price change function:  $EPri = EPri + 5$
- (2) Total quantity of raw materials purchased by the enterprise:  $Q_e = Q_m + Q_{fe}$

- (3) Power generation of the enterprise:  $q = 800 \times Q_e$
- (4) Power generation cost:  $EC_1 = 0.292 \times q$
- (5) Purchase cost of the enterprise:  $EC_2 = EPri \times Q_e$
- (6) The profit of the enterprise:  $W_4 = 0.9 \times \text{Price} \times q - EC_1 - EC_2$

In (6), 0.9 means that 90% of the electricity of the enterprise is for sale and 10% for self-use.

## 5. Analysis of Simulation Results

**5.1. Simulation Interface and Model Verification.** The initial parameters used in the simulation system come from the survey data of farmers in many towns in Inner Mongolia and

Ningxia. Except for unit pruning subsidy, other parameters are assigned in the form of sliding strip. The initial values of number of farmers, number of middlemen, middlemen purchase price, and feed-in tariff are all taken from the survey data as described in Section 4.2.1. The unit pruning cost of farmers usually ranges from 70 to 200 yuan/ton, and the unit processing cost of middlemen ranges from 20 to 35 yuan/ton. The values with high frequency in the survey data are taken as the initial parameter values, which are 100 yuan/ton and 25 yuan/ton, respectively. According to the actual situation of the survey, the initial value of unit pruning subsidy is set to 10.

First, input the initial parameters into the system, as shown in Figure 5.

After inputting the initial parameters, design the view interface of the FBPG raw materials supply chain network, as shown in Figure 6 (the figure shows the transportation of raw materials in a cycle). The gray area is the woodland after pruning ( $n=0$ ), the light green area is the woodland within three years after pruning ( $n<3$ ), and the dark green area represents the woodland that can be used for pruning ( $3\leq n\leq 5$ ). Farmers are represented by small triangles, middlemen are represented by white squares, and the enterprise is represented by a red dot in the center. With the dynamic change of cycle, the interface can visually show the change of pruning and the transportation of raw materials.

Under the setting conditions in Figure 5, export the data obtained from the system to EXCEL and draw the relationship between purchase price and purchase quantity, as shown in Figure 7. According to the above analysis, the amount of raw materials required by the enterprise for power generation is about 100000 tons. If the raw materials purchased by the enterprise can meet the demand when reaching production capacity, the purchase price needs to be between 425 and 435 yuan. At this time, the purchase quantity and purchase price of the enterprise tend to be stable.

Under the setting conditions, the profit of enterprise is negative. Similarly, export the relevant data of enterprise's profit, as shown in Figure 8. After the purchase quantity and purchase price of the enterprise tend to be stable, the loss is about 1.16~1.3 million yuan. In order to make the enterprise more profitable, government subsidy should be gradually increased; that is, the feed-in tariff should be increased.

Therefore, continue to adjust the feed-in tariff to study when enterprise generates profit. When the feed-in tariff reaches 0.85 yuan/kWh, the power generation profit of the enterprise is still negative, with a loss of about 3 million yuan. Continue to gradually increase the feed-in tariff by 0.01 yuan until 0.95 yuan/kWh, and the enterprise can achieve profitability under the condition of meeting the raw materials demand, as shown in Figure 9.

It is not difficult to find that the simulation result is basically consistent with the survey data of existing policy. At present, the purchase price of forest biomass raw materials is 400 yuan per ton, and the enterprise can only purchase 80000~90000 tons of raw materials. In the survey, it is found that most biomass power generation enterprises are in a state of loss at the price of 0.75 yuan/kWh. Therefore, it can

be verified that the model construction is basically reasonable.

*5.2. Simulation Experiments.* The previous section has verified the rationality of the model. Therefore, adjust the parameters from the perspective of supply chain, internal agent, and external environment, and further discuss the influence of parameter change on supply chain. Several comparative experiments are carried out based on the initial parameters.

*5.2.1. From the Perspective of Supply Chain.* The policy subsidy sharing contract can help the whole supply chain achieve coordination [21]. The supply of middlemen is affected by the raw materials market. When the market supply is less than demand, middlemen need to spend a higher price to buy raw materials from other regions. It is difficult for the whole supply chain to achieve coordination in this condition. Middlemen are more inclined to seek a smaller order quantity to ensure that their profit is as large as possible. Therefore, based on the subsidy sharing contract, middlemen are willing to sign a higher order quantity with the enterprise, which improves the enthusiasm for purchase. Similarly, this kind of contract coordination is applicable to farmers and enterprises. Therefore, the enterprise can increase the overall profit of the supply chain by increasing the purchase price. How to determine the optimal subsidy proportion of the enterprise ( $\beta$ ) is very important. In the above initial environment, compare the profit of middleman, enterprise, and farmer, as shown in Figure 10.

When the feed-in tariff remains unchanged (price = 0.95), two periodic nodes are selected, which are the node that does not reach equilibrium and the node that reaches equilibrium. For example, at the two nodes of  $t=7$  (not reaching equilibrium) and  $t=17$  (reaching equilibrium), the two nodes' raw materials purchase price of the enterprise is different, which are 380 yuan per ton and 420 yuan per ton, respectively. Compare the profit of each agent and the whole of the supply chain under the two nodes, and sort out the relevant data of the nodes in Table 7.

When the purchase price of the enterprise increases, the overall profit of the supply chain increases significantly, which verifies the rationality of the experiment. In this experiment, the government subsidy ( $t$ ) is 0.673 yuan/kWh (0.95~0.277), which is equivalent to 538.5 yuan/ton of forest biomass raw materials. At this time, the purchase price is 420 yuan/ton, and the subsidy coefficient ( $\beta$ ) can be calculated as 0.074.

*5.2.2. From the Perspective of the Internal Agent.* The purchase price of enterprise plays a leading role in the whole supply chain. In addition, other parameters also have an impact on the supply of forest biomass raw materials. Such as the unit operating cost of power generation, the unit pruning cost of farmers, and the initial processing cost of middlemen. Reducing these costs can effectively improve the operation efficiency of the supply chain. Next, the paper will

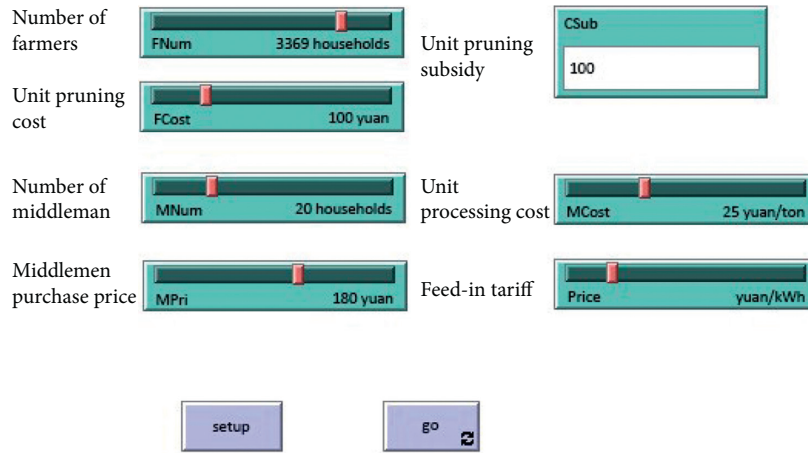


FIGURE 5: Input parameters interface.

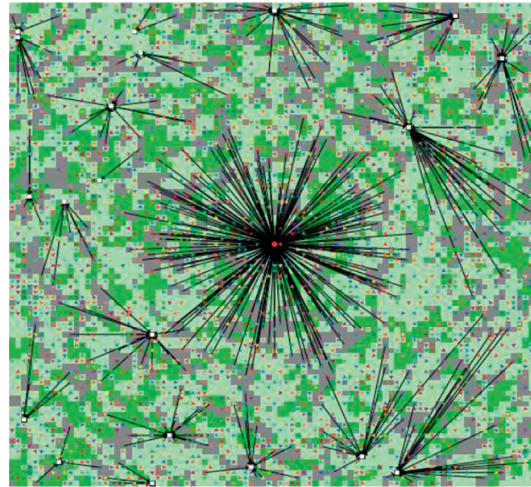


FIGURE 6: Main view interface.

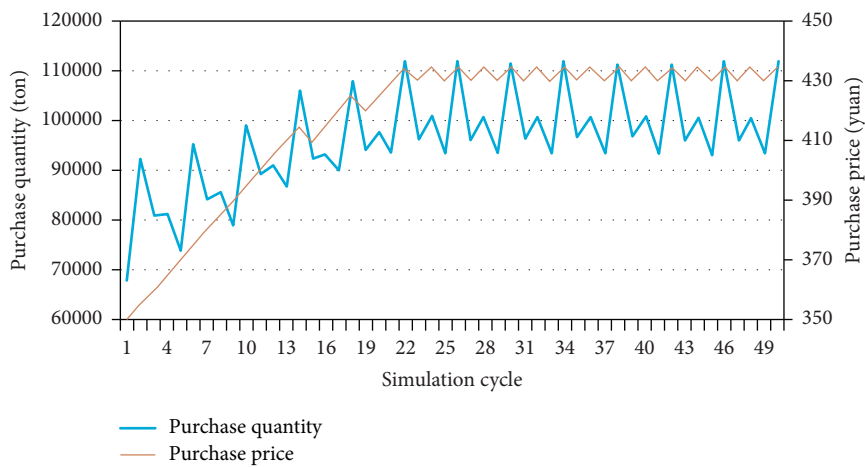


FIGURE 7: Relationship between purchase price and purchase quantity.

study which cost reduction is the most effective in improving the efficiency of forest biomass supply chain. Under the condition that the external environment parameters remain unchanged, several experiments are

carried out, and the results are shown in Table 8. Because the experiments have certain randomness, the results are given according to the average level of many simulation experiments.

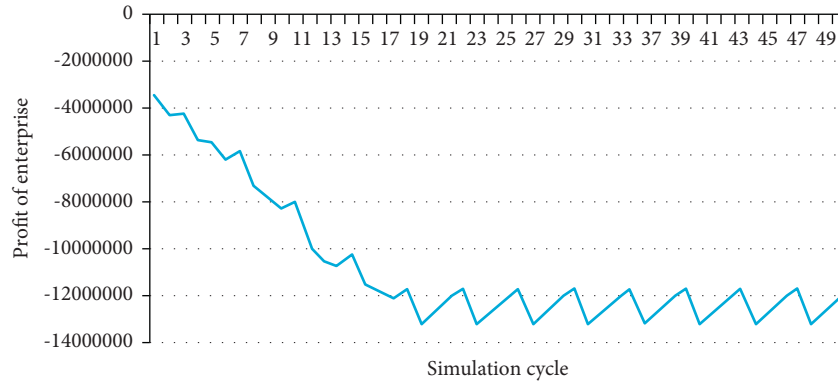


FIGURE 8: Profit of the enterprise (price = 0.75).

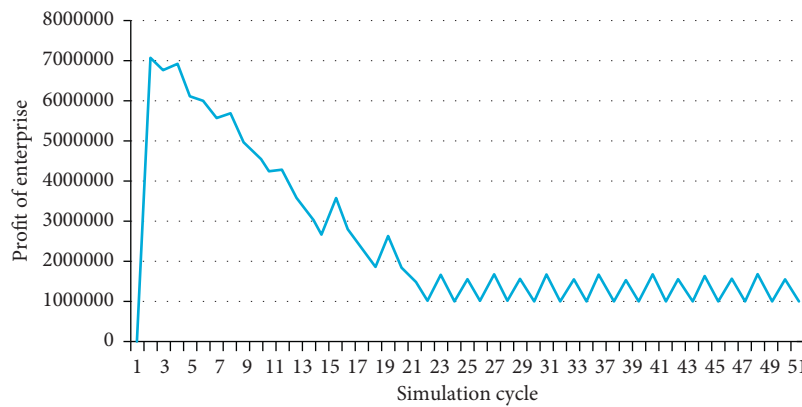


FIGURE 9: Profit of the enterprise (price = 0.95).

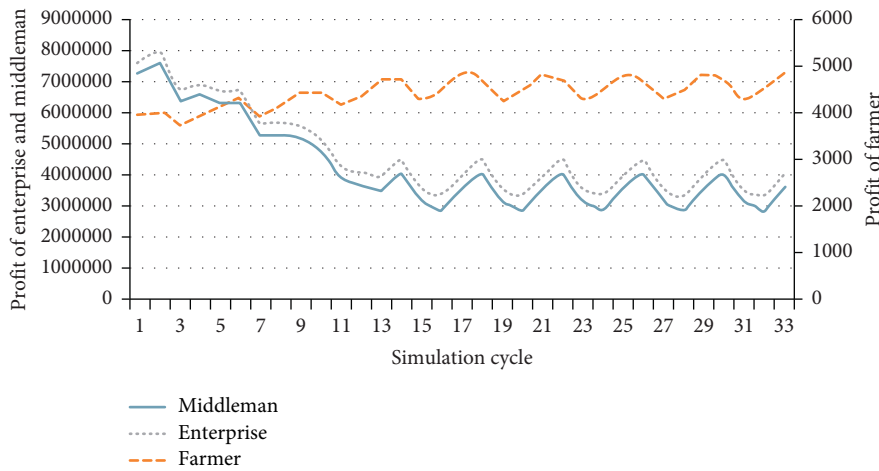


FIGURE 10: The average profit of the agents of supply chain.

The unit pruning cost can continue to be reduced, and the supply chain is more sensitive to the unit pruning cost. Reducing the pruning cost by 30 yuan per mu can greatly reduce the purchase price of the enterprise. Therefore, for the western region in China, improving the mechanization level of raw materials collection and reducing labor cost will be an effective way to improve the efficiency of biomass supply chain. In addition, the reduction of the initial processing cost of middlemen did not cause great changes

in the purchase price of enterprise. Because middlemen have original capital and initial investment, there is little space for the decline of unit initial processing cost. Finally, for enterprises, reducing the operation cost of power generation can effectively avoid the loss. Through technical improvement, strengthening the construction of the management team, and reducing management cost, enterprises can seek survival space under the existing feed-in tariff level.

TABLE 7: Profit index of agents under different purchase prices.

Profit index of agents	Epri = 380	Epri = 420
Number of farmers selling raw materials	331	411
Average profit of farmers	3930	4838
Number of middlemen	20	20
Average profit of middlemen	345867	435437
Number of enterprises	1	1
Average profit of enterprise	5264989	3586342
Total profit of supply chain	13483159	14283500

TABLE 8: Simulation experiment results of changing internal parameters.

Unit pruning cost (farmers)	The initial processing cost (middlemen)	Unit operating cost of power generation (enterprise)	Purchase price of enterprise when reaching production capacity	The feed-in tariff*
100	25	0.292	420–430	0.95
70	25	0.292	285–295	0.65
100	20	0.292	420–430	0.95
100	25	0.25	420–430	0.70

\*The feed-in tariff here refers to the minimum price that can ensure that power generation enterprise does not suffer losses.

### 5.2.3. From the Perspective of the External Environment.

Although the government is not a main agent in the forest biomass supply chain, as an important force to promote the development of the FBPG industry, policy subsidy is an important way to improve the supply of forest biomass raw materials. In addition to the feed-in tariff subsidy mentioned above, the existing subsidies include “immediate levying and refunding” for enterprise tax and pruning subsidy for farmers. In particular, the pruning subsidy of farmers is still in the exploratory stage and has important research value. In addition, it is found in the simulation that the number of farmers and middlemen in the region has a great impact on the whole supply chain. Under the condition that the internal main parameters remain unchanged, several experiments are carried out, and the results are shown in Table 9.

Pruning subsidy can effectively help reduce the purchase price of enterprise and alleviate the pressure on enterprise to purchase raw materials. The number of farmers (or woodland area) in the region has a great impact on the whole supply chain. Therefore, when selecting the site, it is necessary for enterprises to investigate the local forest biomass energy potential and choose the area with rich resources as the power generation address. The number of middlemen will also have a certain impact on the raw materials purchase of enterprises. If the number of middlemen in the region is too small, it will also increase the difficulty for enterprises to acquire raw materials. In particular, it is difficult to acquire raw materials from farmers in remote areas and small areas.

5.3. *Result Analysis.* The simulation experiment is carried out by adjusting the parameter value, and the results are summarized as follows:

- (1) Combined with the actual operation of the current enterprises in China, the current feed-in tariff of 0.75

yuan per kWh (including subsidy) is low for biomass power generation enterprises, and most enterprises are facing the crisis of bankruptcy. The survey also found that 69 farmers believe that the current purchase price of raw materials is unreasonable. The simulation results show that when the feed-in tariff is increased to about 0.95~1.0 yuan per kWh, the biomass power generation enterprises will not suffer losses when reaching the production capacity. At the same time, the purchase price of the enterprises needs to be improved. Under the current purchase price, the supply of raw materials is insufficient. It is more appropriate that the purchase price of raw materials should increase about 25 yuan at the current level of 400 yuan per ton.

- (2) Unit pruning cost has a great impact on the whole supply chain. The increase of pruning cost will lead enterprises to reach the production capacity at a higher purchase price. At the same time, the feed-in tariff that meets the profitable conditions of enterprises will also increase. Therefore, combined with the relevant policies of energy subsidy in the western region of China, the government can adopt appropriate subsidy for purchasing agricultural machinery to realize large-scale mechanized collection and greatly reduce the cost of artificial pruning. The government gives farmers a pruning subsidy, which can also promote some farmers to participate in pruning and increase the supply of raw materials.
- (3) When selecting the site, biomass power generation enterprises need to take the external environment factors into account to ensure that the number of local farmers can provide enough forest biomass raw materials. In addition, the enterprises should establish a long-term and stable partnership with

TABLE 9: Simulation experiment results of changing external parameters.

Number of farmers	Number of middlemen	Pruning subsidy	Purchase price of enterprise when reaching production capacity	The feed-in tariff*
3369	20	0	420–430	0.95
3369	20	100	270–280	0.72
3000	20	0	515–520	1.05
3369	10	0	490–500	1.02

\*The feed-in tariff here refers to the minimum price that can ensure that power generation enterprise does not suffer losses.

middlemen, reasonably plan the number of middlemen, and ensure a stable supply of raw materials.

## 6. Conclusions

FBPG industry is of great significance to protect the ecological environment and promote the increase of farmers' income in the western region of China. Research and survey found that the current bottleneck of industrial development in China is mainly due to the insufficient supply of raw materials and serious losses of enterprises. In order to solve this problem, this paper discusses the coordination of the FBPG supply chain on the premise of uncertain supply. This is a supplement to a subdivided field of straw biomass power generation supply chain research. At the same time, the combination of empirical research and case analysis makes up for the defect of excessive parameter value deviation in the simulation processes and enriches the research content of multiagent simulation of the supply chain. The conclusions of this paper are as follows.

Firstly, it is relatively difficult to collect forest biomass raw materials in China, which has a great relationship with the terrain. Farmers' collection cost is high, and they generally believe that the purchase price of enterprises is low. Because the labor income level has been increasing in recent years, this leads to the low willingness of farmers to collect and sell raw materials. In addition, forest biomass raw materials also have many other values besides being sold to enterprises, such as covering sand layer (sand-prevention) and self-use as an alternative fuel. There are many methods to sell biomass raw materials collected by farmers. The choice of farmers depends on which method can obtain the greatest profit. Therefore, through reasonably increasing the purchase price and giving subsidy for purchasing agricultural machinery, the participation of forest farmers can be improved. It is conducive to the stable supply of raw materials and helps the sustainable and stable development of enterprises.

Secondly, middlemen are an important and stable source for FBPG enterprises to purchase raw materials. At present, the government only provides a feed-in tariff subsidy for biomass power generation enterprises. If the enterprise transfers part of the subsidy to middlemen, the purchase quantity of middlemen will increase, and the profit of the enterprise will also increase. Therefore, enterprises can improve the supply enthusiasm of middlemen by means of increasing purchase price and ensure that the raw materials of FBPG enterprises meet the production capacity demand as much as possible. It also avoids the loss of stock shortage.

Finally, at the current stage, if the enterprise is expected to meet the production capacity demand, the purchase price of 425 yuan per ton is reasonable. In this case, the feed-in tariff can be increased to about 0.95 yuan per kWh to ensure the sustainable operation of biomass power generation enterprises. In addition, reducing the power generation operation cost of enterprises, promoting the use of brush cutters, reducing labor cost, and providing pruning subsidy can also promote the overall efficiency of the supply chain.

Although this paper simulates the reality of the FBPG supply chain in China as much as possible, there are still some details to be further considered due to the complexity of reality. For example, there is a time gap (usually one year) in the issuance of policy subsidy, which may lead to insufficient funds for enterprises to purchase biomass raw materials in that year, and many farmers think that it is difficult to prune, which may lead to a lack of labor. So middlemen need to hire labor from other provinces in China, increasing the cost of pruning.

These complex situations may become important factors hindering the development of biomass enterprises in China. Therefore, the prospects are put forward in the follow-up research:

- (1) The simulation model is only applicable to the forest biomass raw materials supply chain in the western region of China, which can be improved and optimized to make it applicable to other regions.
- (2) Continue to enrich the behavior rules of agents, such as the behavior rules of enterprises. The current setting of enterprise is that price is set on a uniform basis. As a result, farmers close to the enterprise can get rich income, while farmers far away can get less income. So the hierarchical pricing strategy can be further considered. There are different purchase prices for farmers at different distances, which can reduce the cost of enterprises.
- (3) The distribution of farmers, middlemen, and enterprise in the model is different from reality, so a regional map can be introduced to optimize the simulation effect.
- (4) Further discuss the problem of labor shortage in the western region of China.

## Data Availability

All data used in this paper are available from the authors upon request.



## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported by the Fundamental Research Funds for the Central Universities (no. 2021SRZ05) and Beijing Social Science Foundation Planning Project (no. 21GLB036).

## References

- [1] L. Zhang and C. Zhang, "Review on the development of forest biomass energy," *Inquiry into Economic Issues*, vol. 10, pp. 186–190, 2012.
- [2] A. X. Geng, W. Q. Pan, and H. Q. Yang, "Quantifying the mitigating effects and benefits from substituting wood biomass for coal in energy production in China," *Resources Science*, vol. 42, no. 3, pp. 536–547, 2020.
- [3] D. Yu, *Evaluation of forest Biomass Energy Resources Supply Capacity and Analysis of Influence Factors*, Beijing Forestry University, Beijing, China, 2016.
- [4] D. Jiang, K. Sun, W. Qi et al., "Development of biomass power generation industry and suggestions in China," *Renewable Energy Resources*, vol. 32, no. 04, pp. 542–546, 2014.
- [5] Z. Tan, K. Chen, and P. Liu, "Possibilities and challenges of China's forestry biomass resource utilization," *Renewable and Sustainable Energy Reviews*, vol. 41, pp. 368–378, 2015.
- [6] F. Lai and F. Mi, "Thoughts on the role and development model of government in forest biomass power generation industry in China," *World Forestry Research*, vol. 33, no. 2, pp. 106–111, 2020.
- [7] L. Luo, "Thoughts on developing the supply of forest biomass energy raw materials in China," *Journal of Shandong Forestry Science and Technology*, vol. 42, no. 6, pp. 101–105, 2012.
- [8] Q. Peng, Y. Guo, C. Wang, Y. Wen, and K. Chen, "The influence of market orientation on the performance of forest biomass resources enterprises," *Forestry Economics*, vol. 41, no. 11, pp. 10–17+33, 2019.
- [9] P. Hanafizadeh and M. H. Sherkat, "Designing fuzzy-genetic learner model based on multi-agent systems in supply chain management," *Expert Systems with Applications*, vol. 36, no. 6, pp. 10120–10134, 2009.
- [10] P. Wang, H. Xiong, X. Zhai, and C. Chen, "Organizational structure evolution of supply chain and quality and safety of agricultural products computer simulation based on Net-Logo," *Journal of Agrotechnical Economics*, vol. 8, pp. 64–72, 2015.
- [11] R. Zhang and J. Lin, "Influence of the Chinese government subsidy policies on supply chain members' profits: an agent-based modeling and simulation approach," *Journal of Advanced Computational Intelligence and Intelligent Informatics*, vol. 20, no. 4, pp. 623–632, 2016.
- [12] L. Terrada and H. Ouajji, "Multi-agents system implementation for supply chain management making-decision," *Procedia Computer Science*, vol. 177, no. C, pp. 624–630, 2020.
- [13] A. Awaga, W. Xu, L. Liu, and Y. Zhang, "Evolutionary game of green manufacturing mode of enterprises under the influence of government reward and punishment," *Advances in Production Engineering & Management*, vol. 14, no. 4, pp. 416–430, 2020.
- [14] Y. Shastri, L. Rodríguez, A. Hansen, and K. C. Ting, "Agent-based analysis of biomass feedstock production dynamics," *Bioenergy Research*, vol. 4, no. 4, pp. 258–275, 2011.
- [15] Q. Li, *Study on the Evolution and Development of Biomass Utilization Industry System*, Tsinghua University, Beijing, China, 2012.
- [16] Y. Luo, *Game Theory Models of the Biomass Supply Chain for Power Generation Integrating Rural Official Organizations*, North China Electric Power University, Beijing, China, 2019.
- [17] Y. Zuo, *The Study on Forest Insurance Participants' Behavior Based on Complex Adaptive System Theory*, Beijing Forestry University, Beijing, China, 2015.
- [18] Z. H. Su, "Three-carbon economy" industrial sand-benefit analysis," *Forestry Economics*, vol. 7, pp. 26-27+67, 2011.
- [19] Q. Y. Wei, *Research on Supply Chain Logistics Cost of Straw for Biomass Power Generation*, China Agricultural University, Beijing, China, 2014.
- [20] Z. Song, Y. Zhu, F. Lu, and M. Dai, "Analysis on raw material collecting radius of forest biomass: a case study on bio-power generation project in Jinggu County," *Forestry Construction*, vol. 6, pp. 66–68, 2014.
- [21] Y. Qi, *Research on Complex System Simulation of Forest Biomass Power Supply Chain*, Beijing Forestry University, Beijing, China, 2018.

## Research Article

# Exploratory Research on Key Technology of Human-Computer Interactive 2.5-Minute Fast Digital Early Warning for Mild Cognitive Impairment

Nan Li,<sup>1</sup> Xiaotong Yang,<sup>2,3</sup> Wencai Du,<sup>4</sup> Atsushi Ogihara,<sup>5</sup> Siyu Zhou,<sup>6</sup> Xiaowen Ma,<sup>2,3</sup> Yujia Wang,<sup>2,3</sup> Shuwu Li,<sup>7</sup> and Kai Li<sup>2,3</sup> 

<sup>1</sup>School of Humanities and Management, Zhejiang Chinese Medical University, Hangzhou 310053, China

<sup>2</sup>School of Medical Technology and Information Engineering, Zhejiang Chinese Medical University, Hangzhou 310053, China

<sup>3</sup>Zhejiang-Japan Joint Laboratory of Digital Diagnosis and Treatment and Equipment for Major Brain Diseases, Zhejiang Chinese Medical University, Hangzhou 310053, China

<sup>4</sup>University of Saint Joseph, Estrada Marginal da Ilha Verde 14-17, Macau, China

<sup>5</sup>Department of Health Sciences and Social Welfare, Faculty of Human Sciences, Waseda University, Tokorozawa, Japan

<sup>6</sup>School of Public Health, Hangzhou Normal University, Hangzhou, China

<sup>7</sup>Kaibo Medical Equipment (Hangzhou) Co., Ltd., Hangzhou 310052, China

Correspondence should be addressed to Kai Li; [kaili@fudan.edu.cn](mailto:kaili@fudan.edu.cn)

Received 13 December 2021; Revised 20 February 2022; Accepted 24 February 2022; Published 29 March 2022

Academic Editor: Daqing Gong

Copyright © 2022 Nan Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Objective.** As the preclinical stage of Alzheimer's disease (AD), Mild Cognitive Impairment (MCI) is characterized by hidden onset, which is difficult to detect early. Traditional neuropsychological scales are main tools used for assessing MCI. However, due to its strong subjectivity and the influence of many factors such as subjects' educational background, language and hearing ability, and time cost, its accuracy as the standard of early screening is low. Therefore, the purpose of this paper is to propose a new key technology of fast digital early warning for MCI based on eye movement objective data analysis. **Methodology.** Firstly, four exploratory indexes (test durations, correlation degree, lengths of gaze trajectory, and drift rate) of MCI early warning are determined based on the relevant literature research and semistructured expert interview; secondly, the eye movement state is captured based on the eye tracker to realize the data extraction of four exploratory indexes. On this basis, the human-computer interactive 2.5-minute fast digital early warning paradigm for MCI is designed; thirdly, the rationality of the four early warning indexes proposed in this paper and their early warning effectiveness on MCI are verified. **Results.** Through the small sample test of human-computer interactive 2.5 fast digital early warning paradigm for MCI conducted by 32 elderly people aged 70–90 in a medical institution in Hangzhou, the two indexes of “correlation degree” and “drift rate” with statistical differences are selected. The experiment results show that AUC of this MCI early warning paradigm is 0.824. **Conclusion.** The key technology of human-computer interactive 2.5 fast digital early warning for MCI proposed in this paper overcomes the limitations of the existing MCI early warning tools, such as low objectification level, high dependence on professional doctors, long test time, requiring high educational level, and so on. The experiment results show that the early warning technology, as a new generation of objective and effective digital early warning tool, can realize 2.5-minute fast and high-precision preliminary screening and early warning for MCI in the elderly.

## 1. Introduction

Mild cognitive impairment (MCI) is an unstable clinical transition state between normal age and dementia. The study

found that people with MCI above 70 have a high prevalence rate, and 21.8% of them progress to dementia 3~5 years after the onset of MCI. The possibility of transferring to Alzheimer's disease (AD) in patients with MCI is 2.8 times

higher than that in healthy aging [1, 2]. As we all know, Alzheimer's disease is still incurable, whose pathological changes are difficult to reverse and the curative effect is poor in the middle and late stage, bringing psychological, physiological, and economic pressure to the patients and their families [3]. However, studies have found that if patients are identified early and reasonably intervened in MCI stage, the reversal rate of mild cognitive impairment to normal cognitive function is more than 30% [4]. Therefore, moving the MCI early warning window forward at the daily home level based on the key technology of human-computer interactive 2.5-minute fast digital early warning is of great significance to improve the life quality of patients and reduce the social burden.

According to the diagnostic guidelines of MCI, its clinical diagnosis is mainly based on medical history collection and neuropsychological scale assessment [5]. Although the neuropsychological scales commonly used in clinic have good reliability and wide application in screening cognitive impairment, some defects are still widespread, such as strong subjectivity, long time consumption, insufficient sensitivity for MCI early warning, and so on [6]. Thus, in order to make up for the deficiencies of cognitive screening scales, more and more computerized cognitive assessment tools [7, 8] appear, but most of them need professionals to carry out the standardized operation, and the subjects need to go to the clinic or examination centers, which immensely limit the universality of detection.

Moreover, the latest research found that, compared with other cognitive fields, the earliest occurrence of cognitive impairment in MCI patients is the decline of visuospatial function, and at present, many studies have used eye movement related examinations to assess early cognitive impairment [6, 9, 10]. At the same time, the powerful computing and data processing ability of intelligent algorithm has its unique advantages in the analysis of eye movement examination results and eye movement trajectory. Many studies have shown that task assessment saccade mode based on eye tracking can be used to detect MCI and early signs of cognitive decline. Detection indexes and forms vary from research to research [11, 12], but all show that task paradigm can be designed to check various visual features of saccade or eyes under relevant stimuli, and eye tracker can be used for activity monitoring to realize extensive and convenient cognitive function examinations.

Based on the summary of previous studies, it can be found that the current tools for MCI early warning and screening have various shortcomings at different levels. Most of scales and questionnaires which are the most commonly used have deficiencies such as long time-consuming and lack of objective assessment data, while the computerized cognitive assessment tools developed in the early stage also have the characteristics such as various inspection types and parameters and complex calculation process and large population variability.

Thus, taking advantage of noninvasive, short time consumption, low cost, and high feasibility of eye movement tracking data acquisition, this paper puts forward four exploratory MCI early warning indexes including "test

duration," "lengths of gaze trajectory," "correlation degree," and "drift rate" based on the review and summary of relevant research literature and semistructured expert interviews, in visuospatial function and dynamic eye movement monitoring perspectives.

At the same time, based on the key technology of digital early warning, a human-computer interactive 2.5-minute fast digital early warning paradigm for MCI is designed. Finally, the elderly aged 70–90 with clinical diagnosis conclusion of MCI are included in the subjects, and the human-computer interactive 2.5-minute fast digital early warning paradigm designed in this paper is benchmarked with the clinical diagnosis conclusion, so as to verify the effectiveness of the new digital early warning paradigm in this study.

The new early warning key technology studied in this paper has the advantages of short time consumption and simple operation together with objective and accurate assessment, which can realize the real-time feedback of eye movement trajectory and the accurate preliminary early warning for MCI in the elderly population in 2.5 minutes.

The rest of the paper is organized as follows. In Section 2, this paper introduces the related work on MCI early warning in the literature. In Section 3, this paper explains the research methods of this paper. And this paper introduces the relevant work of the experiment, including the paradigm and principle of experiment, the key technologies of digital early warning, and basic information of experiment in Section 4. In Section 5, this paper analyzes the data processing results. A short discussion of the proposed scheme is presented in Section 6.

## 2. Related Work

By reviewing the literature on MCI early warning methods in Pubmed, Web of Science, and other literature libraries in recent 10 years, we find that, compared with other cognitive fields, visuospatial dysfunction may be the earliest cognitive impairment field in MCI patients [13, 14]. A research [15] showed that there is no significant difference between AD mice and normal control mice when performing olfactory tasks, but on the contrary, the visuospatial function including percentage of visuospatial correct decisions and speed of visuospatial relearning of AD mice was significantly worse than that of control mice, which also confirmed that the decline of visuospatial function may be used as an early sensitive index for MCI screening. A study [16] shows that the scores of spatial structure and connection test of MCI patients are low, speculating that visual reasoning, connection, and spatial structure test may be sensitive to the early diagnosis of MCI patients, or the cognitive function reflected by such test, that is, visuospatial and executive function, has diagnostic effect for MCI early diagnosis. In addition, some pathologists [17] found that the initial lesion location of numerous patients with MCI is not the well-known hippocampal region, but the visual contact cortex, which also shows that the loss of visuospatial function may occur before the symptoms of hippocampal involvement (such as memory decline).

At present, many studies have used examinations related to eye movement to assess early cognitive impairment. Study of Bylsma has found that frequency of saccadic intrusion of patients with cognitive impairment may increase due to the decline of attention and working memory function, which is negatively correlated with MMSE scores [18]. VisMET, designed and developed by Haque et al. [19], can assess the visuospatial memory of healthy aging and mild cognitive impairment by analyzing the subjects' gaze points and gaze duration, so as to realize assessment for MCI.

With the continuous development of modern medicine, there are more and more types of computerized cognitive screening systems. Bartoli et al. [20] used Omni robot to control and move the patient's arm to complete a test and analyzed the visuospatial ability of MCI patients through the recorded trajectory and eye movement data. Nevertheless, this study analyzed many variables and needs long time to complete, so it is difficult for the elderly. A computerized cognitive screening test developed by Canadian researchers can screen cognitive impairment through three cognitive domains: information processing speed, working memory, and executive function [21], but it takes 10 minutes and lacks scalability for rapid screening of large-scale populations. Besides, Papp et al. team [22] developed a long-term episodic memory management and measurement method suitable for those at risk of AD by using image stimulation from the perspective of episodic memory. Subjects can detect episodic memory ability through a mobile application, but it mainly assesses cognitive ability from the memory dimension, with low efficiency of MCI early screening and warning.

Consequently, although the existing cognitive screening information system has many advantages, the screening for MCI patients still has the following deficiencies:

- (1) Systematic screening needs professional doctors. At present, it is difficult to carry out large-scale community screening under the current situation of lacking doctor resources.
- (2) The system test is highly difficult and time is too long, so it is not suitable for fast primary screening in large-area home crowd.
- (3) The design of test paradigm is not objective enough, the degree of standardization is low, and there is a lack of objective and quantitative assessment indexes. The early warning effectiveness is not enough, and it is difficult to meet the needs of screening.
- (4) There are still few studies on digital early warning screening for detecting MCI from the perspective of visuospatial dysfunction.

See Table 1 for more detailed analysis of deficiencies, which also shows the research gaps found in previous studies in the field of MCI early warning.

At the same time, we found that numerous studies assess early cognitive impairment through eye movement data, indicating that eye movement test is a relatively simple and objective assessment and screening tool for MCI at present.

Although detection indexes and forms vary from research to research, the assessment of eye movement test in these studies often focuses on test durations, eye drift, and so on. Based on this, aiming at the shortcomings of MCI early warning at present, a human-computer interactive digital early warning paradigm for MCI is designed based on visuospatial function assessment and dynamic eye movement monitoring. It realizes eye movement dynamic follow-up and real-time feedback under time sequence and outputs visual dynamic analysis results in time on the basis of the concept of scene activation. Concurrently, the fast, accurate, and digital early warning and screening for MCI high-risk population based on eye movement objective data analysis is realized.

### 3. Methodology

*3.1. Semistructured Interview.* In order to get more reliable information, we use semistructured interview to determine the effective early warning indexes for MCI. The 20 experts participating in the interview are composed of neurologists, interdisciplinary scholars of neurocomputing, and elderly health management experts, who all have rich experience in the assessment and diagnosis of MCI in large general hospitals.

Based on previous literature review [25–28], this paper formulates an interview outline. By designing the outline of behavioural event interview, the interviewees are reviewed in an open-ended manner so as to dig out the indexes related to MCI early warning. Three days before the formal interview, the interview outline was provided for the interviewees for preparation, so as to improve the effect and quality of the interview. In addition, before the formal implementation of the interview, this paper clearly informed the interview process and the purpose of the obtained data, so as to avoid the interviewees' worries. The outline of the interview is shown in Table 2.

*3.2. Key Technology of Eye Movement Objective Data Acquisition.* The objective data acquisition of eye movement is realized by Tobii eye tracker, composed of two eye movement sensors, dark pupil illumination light sources, bright pupil illumination light source, and multiple signal processing chips, whose internal structure is shown in Figure 1.

Tobii eye tracker uses multiple near-infrared light sources as reference points for auxiliary analysis. By collecting the reflected light from the pupil and cornea, the eye tracker analyzes the relative position of eyes and then obtains the focus of the user's line of sight. On the basis of multi-reference point complementary technology, the user's head trajectory compensation is realized to ensure the accuracy of the collected data. Figure 2 shows light source reflection diagram. The technology adopted by Tobii eye tracker is an improved version of the traditional pupil center cornea reflection (PCCR) eye tracking technology (US patent us7572008) [29], and the principle is as follows:

TABLE 1: Previous studies and research gap.

Research	Method	Limitation
P. Maruff, Y. Y Lim, D. Darby, et al. [23]	Novel method for rapid assessment of cognitive impairment using high-performance eye tracking technology	To assess by staring at still images, with long time consumption, not activating the human-computer interactive scene, lacking dynamic follow-up of eye movement in real time under time sequence, not realizing the automatic test function
A. Oyama, S. Takeda, Y. Ito, T. Nakajima, et al. [24]	Through the functional assessment from visuospatial memory ability realizing MCI early warning	
R. U. Haque, C. M. Manzanaraes, L. N. Brown, et al. [19]	Through the assessment from visuospatial learning ability and memory defect for MCI early warning	
E. Bartoli, F. Caso, G. Magnani, and G. Baud-Bovy [20].	Using robot to control and move the patient's arm to complete a test and analyzing the visuospatial ability of MCI patients through the recorded trajectory and eye movement data	The operation is difficult for the elderly and takes long time
I. E. Plattner, L. Mbakile-Mahlanza, S. Marobela, et al [21]	Screening cognitive impairment through three cognitive domains: information processing speed, working memory, and executive function	Long time consumption
R. F. Buckley, K. P. Sparks, K. V. Papp [7]	Using image stimulation from the perspective of episodic memory for nonclinical AD	The main function is to assist in the diagnosis and judge the progress and severity of AD, and the early warning efficiency of MCI is low

$$\begin{aligned}
d &\sim \theta \\
h &\approx x\theta \\
h &= 2r \sin \frac{\varphi}{2} \\
\varphi &= \varphi' - \frac{\theta}{2} \approx \varphi' \\
\varphi' &\approx \arctan \frac{a}{x} \\
d &\sim \frac{r}{x} \sin \left[ \arctan \left( \frac{a}{x} \right) \right]
\end{aligned} \tag{1}$$

Based on the principle shown in Figure 3, the human-computer interactive 2.5-minute fast digital early warning paradigm for MCI studied in this paper is completed on a computer equipped with eye tracker. Below the computer display is the eye tracker used in this study. In order to explain the calculation method more clearly, we define the position information of the subject's line of sight on the display collected by the eye tracker as "gaze sample," whose sampling rate studied in this paper is 60 Hz, which is about to collect the subject's "gaze sample" every 0.0167 seconds. The principle of human-computer interactive eye tracking is shown in Figure 3.

## 4. Experimental Design

*4.1. Experimental Design of Human-Computer Interactive Fast Digital Early Warning Paradigm for MCI.* The key technology of early warning proposed in this paper serves the elderly high-risk groups with cognitive impairment and is committed to the early warning and screening of MCI in the home or community scene, which should meet the use needs of high-risk groups with cognitive impairment. First of all, designing digital games for elderly users need to be user-centered, which should be suitable for the knowledge and understanding of elderly players [30]. Therefore, this

study is based on the principle of human factors engineering, follows the user-centered design principle, and finally completes the game design based on the preexperiment of multiple schemes in the early stage.

Through the preliminary preexperiment, we found that most of the elderly in the community have varying degrees of vision problems, and most of them are right-handed. Hence, considering the actual situation of users and the reasonable distance of data measurement, this study uses a 32-inch 4K ultraclear display screen, equipped with high-performance miniaturized eye tracker (Tobii eye tracker 5) and Xbox handle as operating device. The eye tracker can be installed under the computer display for eye movement perception and acquisition. During the experiment, each subject needs to sit 55 cm away from the device to meet the requirements of vision and data acquisition.

The subjects complete the test through operating handle, which is easy to move and has a round touch. Besides, the rocker is mainly used in the operation, which is an easier operation method to master, compared with the keyboard and mouse. It is simple to use and can enhance the sense of main control and have a high degree of adaptability to the aging.

In terms of paradigm design, the interface designed in this study is mainly black and white, and the object controlled by the subjects is a red ball. The color contrast between red ball and the interface is strong, making the target prominent. As for designing obstacles and eliminating targets, the subjects are mostly right-handed operation, so the main obstacles are set on the left side of the interface to facilitate the subjects to operate in a relatively fixed posture. Meanwhile, considering that the elderly have obstacles in understanding that the targets need to be removed from the paradigm, the four targets in the interface are set to dynamic 3D rotation. See Figure 4 for overall presentation.

Through the preexperiment of 80 samples in the early stage, it is found that the total time for most people with normal cognition to complete the paradigm will not exceed



TABLE 2: Interview outlines.

Interview outlines
1. What do you think is the goal of MCI early warning?
2. What do you think is the most effective way to assess MCI?
3. Which objective early warning indexes do you think can realize MCI early warning?
4. What do you think are the shortcomings of the current MCI early warning methods?
5. What are your expectations and suggestions for MCI early warning mode in the future?

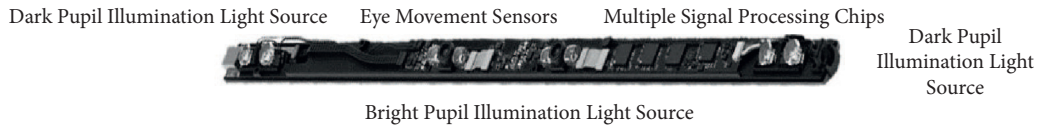


FIGURE 1: Internal structure of eye tracker.

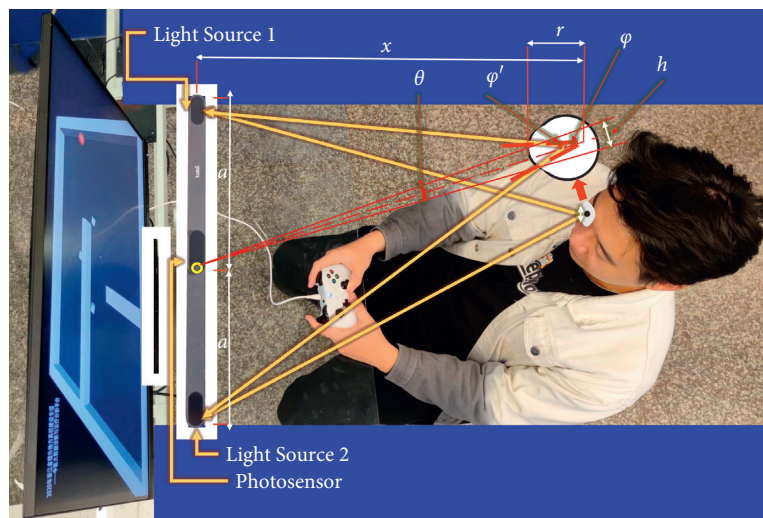


FIGURE 2: Light source reflection diagram.

2.5 minutes (150 seconds). Thus, when designing the paradigm, 2.5 minutes will be taken as cut-off point of the reference paradigm time. Subjects who still fail to complete the paradigm after 2.5 minutes are considered to have a high risk of MCI. Figure 5 shows a case of experiment.

The paradigm we designed consists of three modules: information entry, white cube elimination, and data feedback. In order to ensure that the equipment can accurately detect the subject's eyes, the equipment needs to be calibrated before the test. The test rules and the method of operating the equipment will be orally informed to the subjects by professionals. In the module of information entry, the staff can click the start button to conduct the formal test after entering the subject's name. In the module of white cube elimination, subjects need to use the rocker of the Xbox handle to control the movement of the red ball to collide and eliminate four white cubes. The test paradigm is shown in Figure 6.

After the four white cubes are all eliminated, the subjects can receive their test durations in the module of data feedback on the screen. What is more, the eye movement trajectory will be displayed also. Both the test data are fed back to the user as soon as the test ends, as shown in Figure 7.

#### 4.2. Experiment Implementation

**4.2.1. Data Acquisition and Principle in Test Paradigm.** Based on the summary of relevant literature and semi-structured expert interviews, we dug out four exploratory MCI early warning indexes including test durations, lengths of gaze trajectory, correlation degree, and drift rate. The digital paradigm and mathematical analysis technology are applied to assess the visuospatial function and dynamically monitor the eye movement, so as to realize the preliminary screening and early warning for MCI. This technology is mainly realized by designing the paradigm scene in which the subjects control the red ball to eliminate the four white cubes in the virtual environment through collision. Finally, the test durations and the dual trajectories graph of gaze and read ball will be fed back to users. The framework of key technology of human-computer interactive fast digital early warning for MCI is shown in Figure 8.

Moreover, in order to observe and analyze data more accurately and intuitively, Ecarts is used to visualize the data in this study. As shown in Figure 9, in the 3D coordinate diagram, the  $X$  axis represents the abscissa of the red ball movement trajectory and gaze trajectory, the  $Y$  axis represents the ordinate of the red ball movement trajectory and



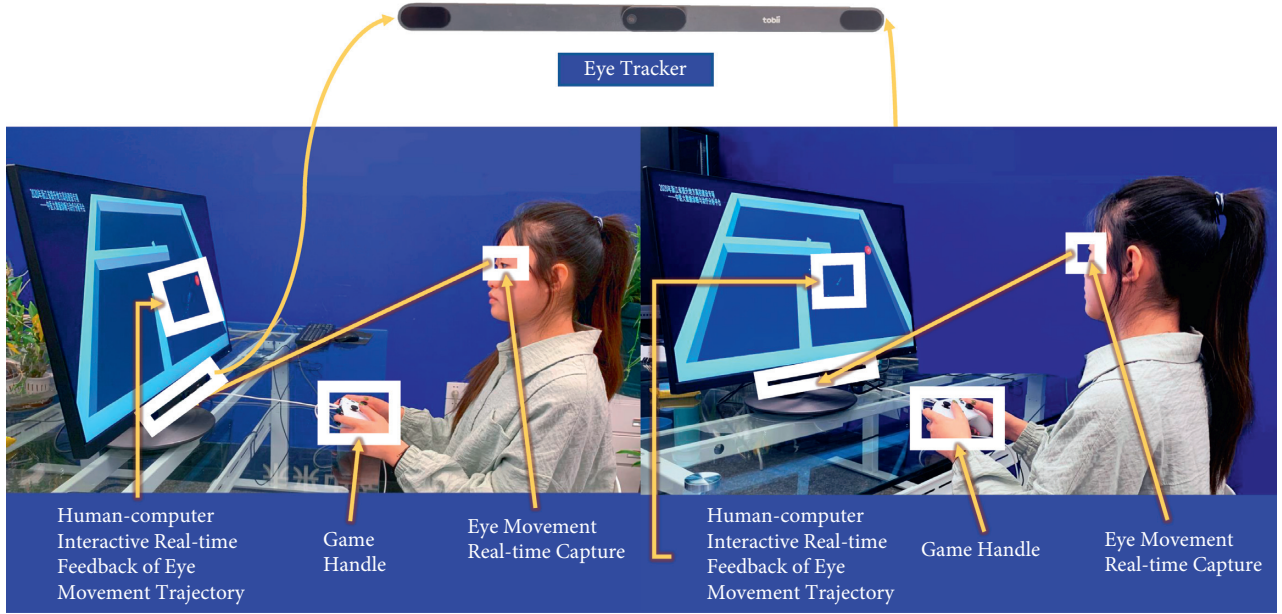


FIGURE 3: The principle of human-computer interactive eye tracking.

gaze trajectory, and the  $Z$  axis represents the time of the paradigm. Then, the gaze sample position, red ball position, and four white cube positions in the paradigm are represented in the figure with different color labels, respectively. In the graph, dark red curve represents the red ball movement trajectory and navy-blue curve represents the gaze sample trajectory. Purple, yellow, green, and red lines, respectively, to represent the positions of four white cubes.

Subjects can obtain the visualization results of paradigm data after the paradigm ends and retain the corresponding trajectory by checking different labels. 3D dual trajectories graph from different angles is shown in Figure 10. Non-technical personnel can make an intuitive comparison by observing the similarity between the gaze trajectory and the red ball trajectory and then get a general judgment, while professional data analysis needs further calculation.

**4.2.2. Definition and Quantitative Analysis of Key Indexes in Test Paradigm.** Based on the previous literature review and semistructured interviews, this study identified four key indexes including test durations, lengths of gaze trajectory, correlation degree, and drift rate. In this paradigm, the total time is  $T(s)$ , the sampling time of the gaze samples and the position of the red ball in  $T$  is  $I$ , the time point of single sampling is  $t_i$ , the single gaze sample position is  $Gaze_i$ , and the single position of red ball is  $Ball_i$  ( $0 < i \leq I, i \in N^*$ ).

$L_{Gaze}$  means length of gaze trajectory in this paradigm, which is the total length of the gaze sample trajectory. The calculation method is as follows:

Make  $d_i$  the distance between two adjacent gaze samples:

$$d_i = \sqrt{(X_{Gaze_{i+1}} - X_{Gaze_i})^2 + (Y_{Gaze_{i+1}} - Y_{Gaze_i})^2} \quad (1 \leq i \leq I - 1, i \in N^*) \quad (2)$$

where  $X_{Gaze_{i+1}}$  refers to the abscissa of the gaze point at the  $i + 1$ th sampling, and  $X_{Gaze_i}$  refers to the abscissa of the gaze sample at the  $i$ th sampling;  $Y_{Gaze_{i+1}}$  refers to the ordinate of the gaze sample at the  $i + 1$ th sampling, and  $Y_{Gaze_i}$  refers to the ordinate of the gaze sample at the  $i$ th sampling.

Lengths of gaze trajectory  $L_{Gaze}$  can be obtained:

$$L_{Gaze} = \sum_{i=1}^{i=I-1} d_i. \quad (3)$$

In this paradigm, the correlation degree is used to describe the similarity between the red ball trajectory and the gaze trajectory. The algorithm of the correlation degree is based on the correlation coefficient, which is the quantity of the linear correlation degree between the research variables. The formula is as follows:

$$\text{corr}(a, b) = \frac{\text{cov}(a, b)}{\sqrt{\text{var}[a] \times \text{var}[b]}} \quad (4)$$

$\text{cov}(a, b)$  is the covariance between  $a$  and  $b$ ,  $\text{var}[a]$  is the variance of  $a$ , and  $\text{var}[b]$  is the variance of  $b$ .

Because the red ball trajectory (the red ball coordinate position sequence,  $Ball$ ) and the gaze trajectory (the gaze coordinate position sequence,  $Gaze$ ) are not simple one-dimensional random variable data, the abscissa and ordinate of the two sequences are separated to form the corresponding sequence separately.  $X_{Ball}$  is the red ball abscissa sequence,  $Y_{Ball}$  is the red ball ordinate sequence,  $X_{Gaze}$  is the abscissa sequence of gaze samples, and  $Y_{Gaze}$  is the ordinate sequence of gaze samples. At this time, correlation degree of  $X_{Ball}$  and  $X_{Gaze}$ , that is,  $X_{corr}$ , and correlation degree of  $Y_{Ball}$  and  $Y_{Gaze}$ , that is,  $Y_{corr}$ , can be calculated, respectively.

Finally, the correlation degree is obtained:

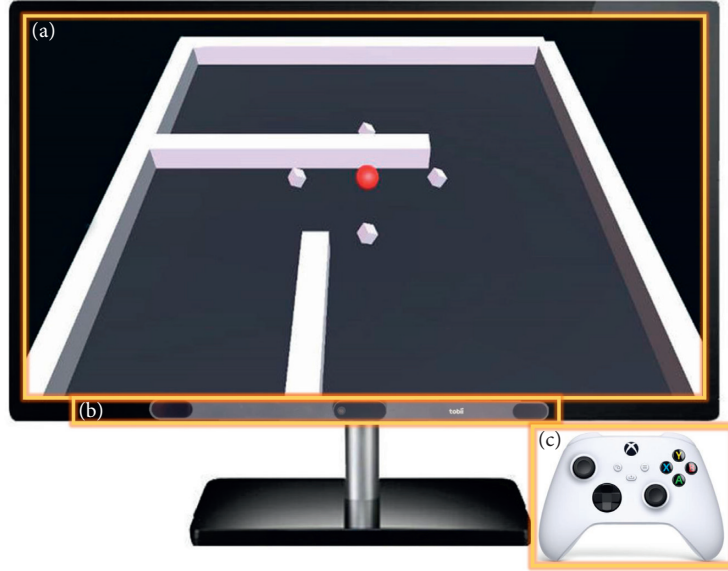


FIGURE 4: Overall presentation. (a) Paradigm interface of red ball. (b) Eye tracker. (c) Handle.

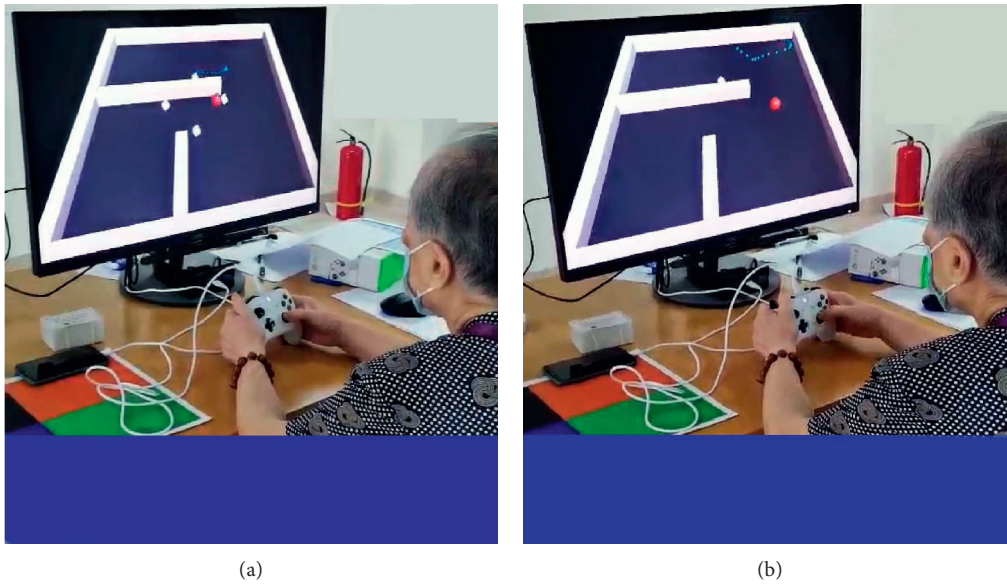


FIGURE 5: A case of experiment. (a) The scene of the test starting. (b) The scene of the test going to end.

$$corr = \frac{X_{corr} + Y_{corr}}{2}. \quad (5)$$

In this study, Ecarts is used to visualize the test results of MCI early warning paradigm, and the 3D dual trajectories graph of the gaze trajectory and red ball trajectory of the subjects is formed. The experimental legend shows that the trajectory of some subjects shows the sharp angle of gaze trajectory and is far away from the red ball trajectory, which is unreasonable, because the goal of the paradigm designed in this paper is to let the subjects control the red ball to eliminate the white cubes. At the same time, when setting the numerical range, this test allows a certain limit of line-of-sight offset caused by vision, sitting position, and instrument

error. Therefore, if the distance between the subject's gaze sample and the red ball position exceeds a certain range at the same time, it is regarded as eye drift.

In this study, the project team explored the value of eye drift threshold between 100 px and 1000 px and took every 50 px as the standard interval, with a total of 21 methods. The Apriori algorithm is used to analyze the correlation between the subjects' gaze sample data under different values and the subjects' final judgment results. The results show that the confidence is the highest at the thresholds of 400 px, 450 px, and 500 px, all reaching 0.89. Furthermore, the project team conducted logistic regression analysis on the three groups of data with the highest confidence. It was found that the prediction effect was the best at the threshold of 450 px.

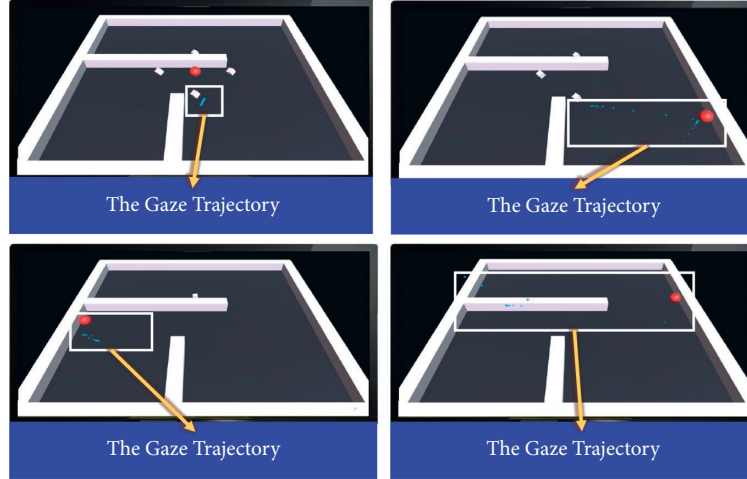


FIGURE 6: Test paradigm for the purpose of eliminating cubes.

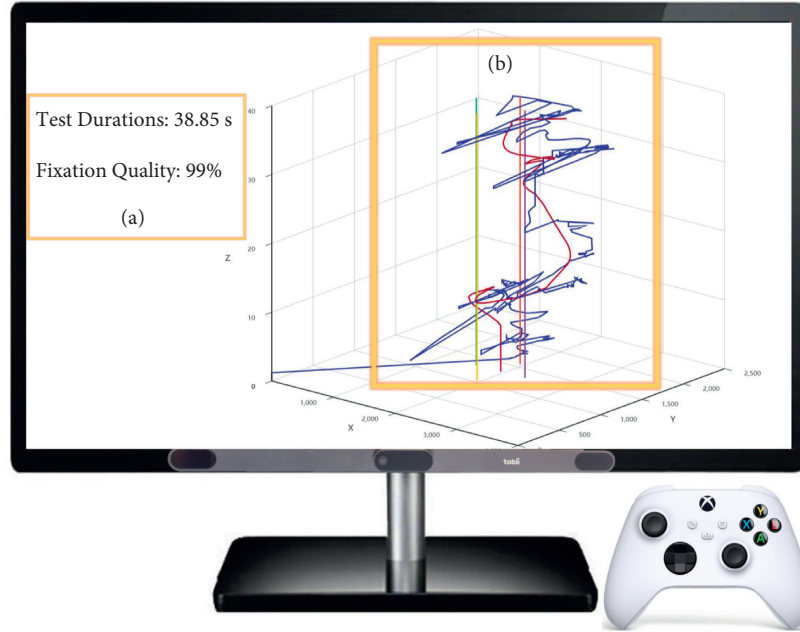


FIGURE 7: Data feedback. (a) Test durations. (b) Gaze trajectory and ball trajectory.

Drift rate is calculated as follows:

In formula (6),  $L_{Gaze-Ball_i}$  is the length between the gaze sample and the position of the red ball at the  $i$ th sampling.

$$L_{Gaze-Ball_i} = \sqrt{(X_{Gaze_i} - X_{Ball_i})^2 + (Y_{Gaze_i} - Y_{Ball_i})^2} \quad (6)$$

Formula (6) is a function to judge whether the gaze drifts away. In formula (7),  $D$  is the number of distances beyond 450 (px) between the red ball and the gaze sample in all samples. That is,  $L_{Gaze-Ball} \geq 450$  (px).

$$f(L_{Gaze-Ball_i}) = \begin{cases} 1 & L_{Gaze-Ball_i} \geq 450 \text{ (px)} \\ 0 & L_{Gaze-Ball_i} < 450 \text{ (px)} \end{cases} \quad (i \leq I, i \in N^*), \quad (7)$$

$$D = \sum_{i=1}^{i=I} f(L_{Gaze-Ball_i}) \quad (0 < i \leq I, i \in N^*) \quad (8)$$

Finally, drift rate is obtained:

$$\text{Drift Rate} = \frac{D}{I} \times 100\% \quad (9)$$

**4.2.3. Participants.** MCI patients and the normal cognitive aged in a medical institution in Hangzhou were taken as the research objects; the human-computer interactive 2.5-minute fast digital early warning paradigm test for MCI was carried out. As four early warning indexes of MCI fast digital early warning, test durations, lengths of gaze trajectory, correlation degree, and drift rate were recorded.

Inclusion criteria of MCI group are as follows: (1) at the age of 70–90 years; (2) with a clear clinical MCI diagnosis; (3) able to complete the test independently. Inclusion criteria of normal cognitive group are as follows: (1)

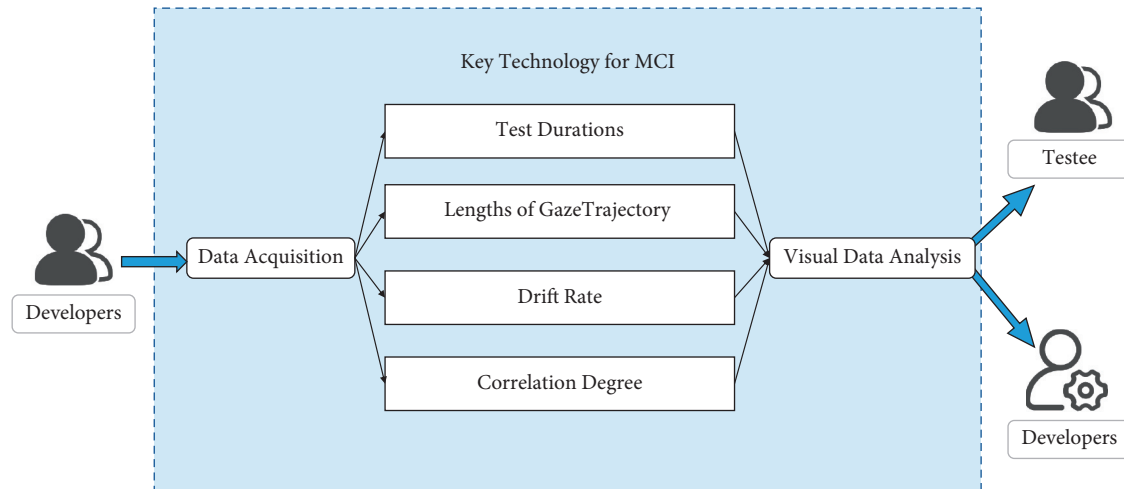


FIGURE 8: Key technology of human-computer interactive fast digital early warning for MCI.

cognitive impairment being excluded clinically; (2) with a clear clinical MCI diagnosis; (3) able to complete the test independently;

See Table 3 for exclusion criteria.

**4.2.4. Rationality Test of Key Indexes for MCI Early Warning.** SPSS 26.0 statistical software is used for data processing. The measurement data conforming to the normal distribution is represented by  $\bar{x} \pm s$  and data between two groups are analyzed by independent sample  $t$ -test. The measurement data not conforming to the normal distribution are represented by the  $M$  (quartile) and data between two groups are analyzed by nonparametric test. Compare the characteristics of test durations, lengths of gaze trajectory, correlation degree, and drift between MCI group and normal cognitive group ( $p < 0.05$ ). The prediction performance for MCI of each index is analyzed by ROC curve and AUC.

## 5. Results

**5.1. Results of Semistructured Interview.** This study summarizes the interview results, as the following shows.

- (1) The important goal of constructing MCI early warning paradigm is to detect and screen MCI quickly and effectively, so as to reduce the conversion rate of MCI to AD and reduce the burden on families and society.
- (2) Eye movement examination is an effective way to detect and assess MCI.
- (3) The objective early warning indexes for MCI mainly include test duration, lengths of gaze trajectory, correlation degree, and drift rate.
- (4) The deficiency of current MCI early warning technology is the lack of objective dynamic assessment index and fast assessment mode.
- (5) It is expected to build a new fast digital early warning model for MCI in the future.

Figure 11 shows indexes and coefficient with high frequency based on semistructured interview.

### 5.2. Results of the Experiment

**5.2.1. Participants Characteristics.** A total of 36 subjects were collected in this experiment, and 4 cases were excluded due to cataract, stroke, and depression. A total of 32 subjects meeting the above criteria were included, including 16 patients with MCI and 16 normal cognitive aged. There was no significant difference in sex ratio ( $p = 0.373$ ), age ( $p = 0.064$ ), and education level ( $p = 0.175$ ) between the two groups, which eliminated the effect of age, gender, and education level on the follow-up test results. See Table 4 for statistical analysis results.

**5.2.2. Comparison of Three Indexes between MCI Group and NC Group.** In order to further determine whether the indexes of key technology of human-computer interactive fast digital early warning between MCI group and NC group is discriminative, this paper statistically analyzed the test durations, lengths of gaze trajectory, correlation degree, and drift rate between two groups. The results are shown in Table 5.

Test durations (s) of MCI group are 50.44 (85.89) and those of NC group are 43.39 (42.22). There is no significant difference in test durations between the two groups ( $p = 0.407$ ). Lengths of gaze trajectory (pixels) in MCI group are 11562.63 (9539.73), and those in NC group are 10474.51 (6638.62). There is no significant difference in the distribution of lengths of gaze trajectory between the two groups ( $p = 0.274$ ). Correlation degrees of MCI group are  $0.74 \pm 0.19$ , and those of NC group are  $0.87 \pm 0.08$ . There is significant difference in correlation degrees between the two groups ( $p = 0.001$ ). Drift rate of MCI group is 0.28 (0.27), and that of NC group is 0.16 (0.18). There is significant difference in drift rate between the two groups ( $p = 0.019$ ). The measured values of correlation degree in MCI group are slightly lower than that of NC group and drift rate of MCI group is greater than that of



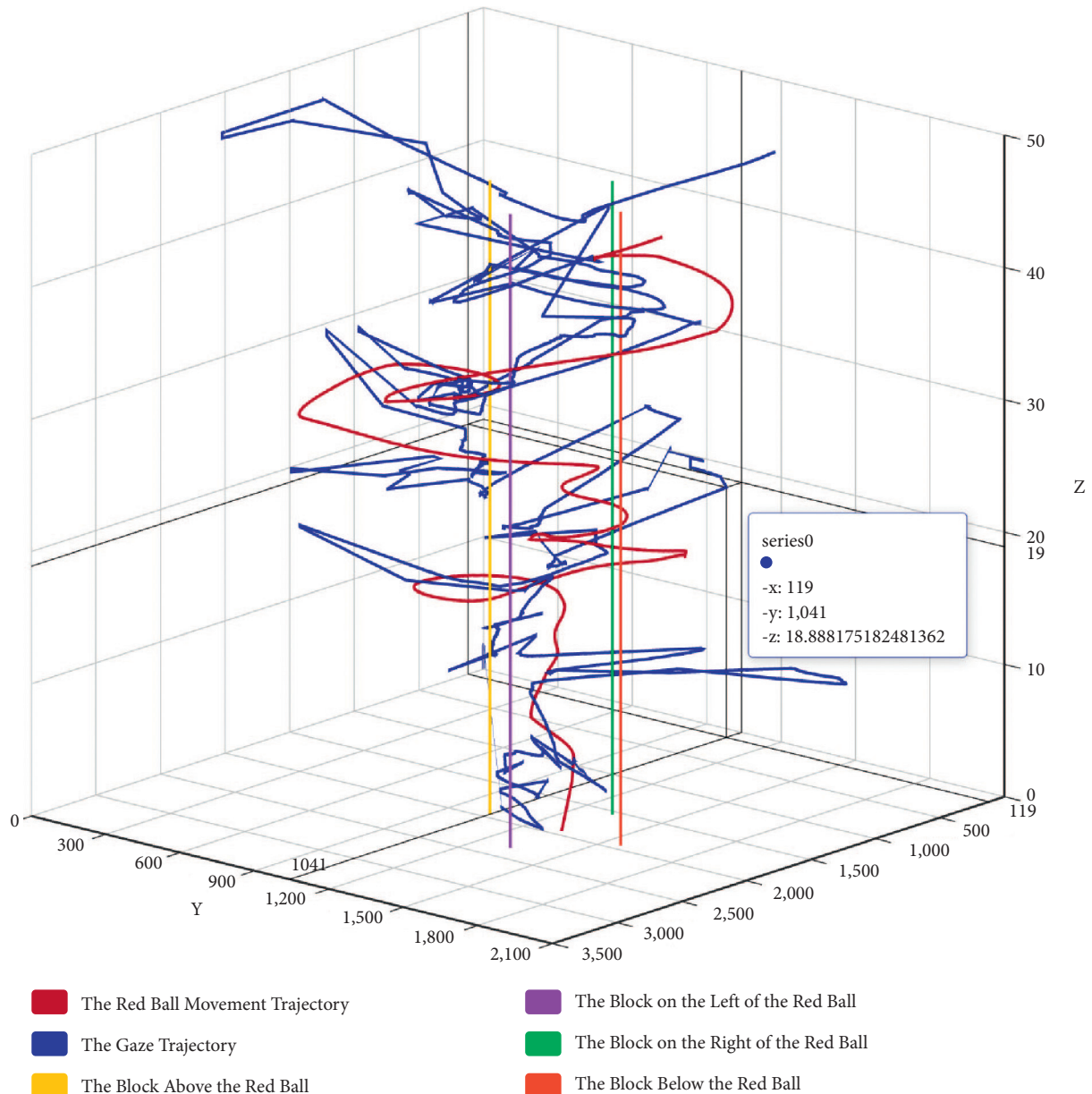


FIGURE 9: A case of 3D dual trajectories graph.

NC group, and the data between the two groups are statistically significant.

Figure 12 is drawn by Graphpad Prism 9.0, showing the general distribution of data between the two groups.

**5.2.3. Assessment Results of Key Indexes Screening Rate Based on ROC Curve.** The prediction performances for MCI early warning of the two indexes (correlation degree and drift rate) in human-computer interactive 2.5-minute fast digital early warning test for MCI are determined using a ROC analysis. An area under the ROC curve (AUC) is used as an index of prediction performance for discriminating MCI from normal cognitive aged.

Firstly, we explored the ROC curve and the area under the ROC curve (AUC) for MCI early warning of correlation

degree and drift rate, respectively. As shown in Figure 13, AUC of correlation degree is 0.809, and AUC of drift rate is 0.742.

In order to further improve the prediction performance for MCI early warning, we paired the two indexes with discriminative significance (correlation degree and drift rate) and operated the ROC curve analysis. As shown in Figure 14, AUC of combining correlation degree and drift rate is 0.824.

## 6. Discussion

Firstly, based on the key technology of human-computer interactive 2.5-minute fast digital early warning, this paper explores the early warning performance of the four indexes of test durations, lengths of gaze trajectory, correlation

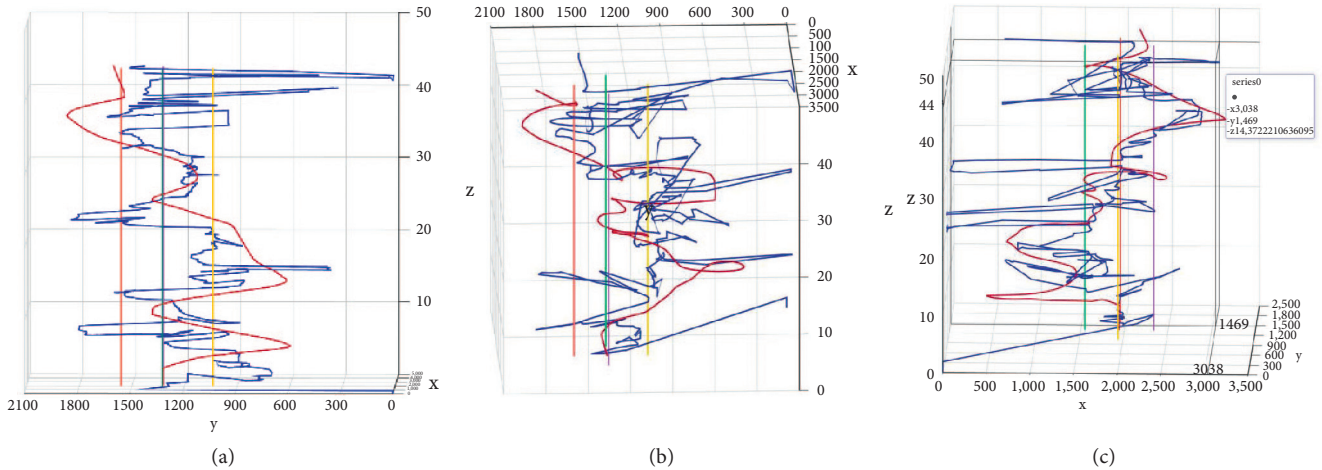


FIGURE 10: 3D dual trajectories graph from different angels. (a), (b), (c) 3D dual trajectories graph from different angels.

TABLE 3: Exclusion criteria.

Serial number	Item
1	Tumour
2	Uncontrollable diabetes
3	Brain trauma
4	History of drug abuse
5	Known learning disabilities
7	Apoplexy
8	Serious medical diseases
9	Depression
10	Nonprimary cognitive impairment
11	Significant visual impairment

degree, and drift rate for MCI through the small sample test of 32 subjects aged 70–90. In the preexperiment, we found that the maximum test durations of the normal cognitive population (including the elderly with unskilled operation) in the human-computer interactive early warning paradigm in this paper are 1 minute and 50 seconds. In order to prevent special situations, we have extended it by 40 seconds (for avoiding some special minor situations during the test, so the elderly with the test durations more than 2.5 minutes can end the test and directly determine being at the risk of MCI. After the experiment, it is found that there is no significant difference in test durations and lengths of gaze trajectory between MCI and NC ( $p > 0.05$ ), and there are significant differences in correlation degree and drift rate between the two groups ( $p < 0.05$ ). The results confirm that the human-computer interactive 2.5-minute fast digital early warning paradigm for MCI designed in this paper is not related to test durations in people with normal cognition and early MCI, if test durations are less than 2.5 minutes, which breaks our view that the longer the test durations are, the more serious the MCI condition is (even when the test duration of MCI patients is less than that of people with normal cognition). Nevertheless, the two MCI early warning indexes, correlation degree and drift rate, still show good early warning effectiveness. Hence, within 2.5 minutes of the early warning prediction paradigm, some normal cognitive aged will increase the test durations due to their unskilled

operation of electronic devices, but the increase will not affect the two key MCI early warning indexes: correlation degree and drift rate.

Secondly, we analyze the early warning performance of correlation degree and drift rate for MCI, respectively. In the analysis of early warning performance of the single index for MCI, the AUC of correlation degree is 0.809; the AUC of drift rate is 0.742, which is consistent with the research conclusion [31] that saccadic intrusions greatly increase in eye movement dynamic examination in patients with impaired cognition. Based on the above results, we can conclude that, under the mode of early warning performance of the single index, the prediction effectiveness of correlation degree is slightly higher than that of drift rate. Thus, the early warning effectiveness of correlation degree is better.

Subsequently, we further explored the ROC curve of MCI early warning after the combination of correlation degree and drift rate. AUC of combining correlation degree and drift rate is 0.824, which is higher than AUC of correlation degree (0.809) and AUC of drift rate (0.742), indicating the MCI prediction performance of combining correlation degree and drift rate is higher than that of single index in the experiment.

Next, in order to further explore the prediction performance of MCI early warning indexes, two indexes (correlation degree and drift rate) with discriminative significance and two key indexes (test durations and lengths of gaze trajectory) are combined for joint early warning analysis. Although there is no significant difference in the distribution of test durations and lengths of gaze trajectory between the two groups in part 5.2.2, the experiment results show that the area under the ROC curve (AUC) is improved when the two key indexes are combined with two indexes with discriminative significance as shown in Figure 15. The area under the ROC curve of combining test durations, lengths of gaze trajectory, correlation degree, and drift rate for MCI early warning and prediction is up to 0.895, which is higher than the early warning effect of the combined use of the two digital biomarkers. The occurrence of these situations may be related to the small sample size of this paper.



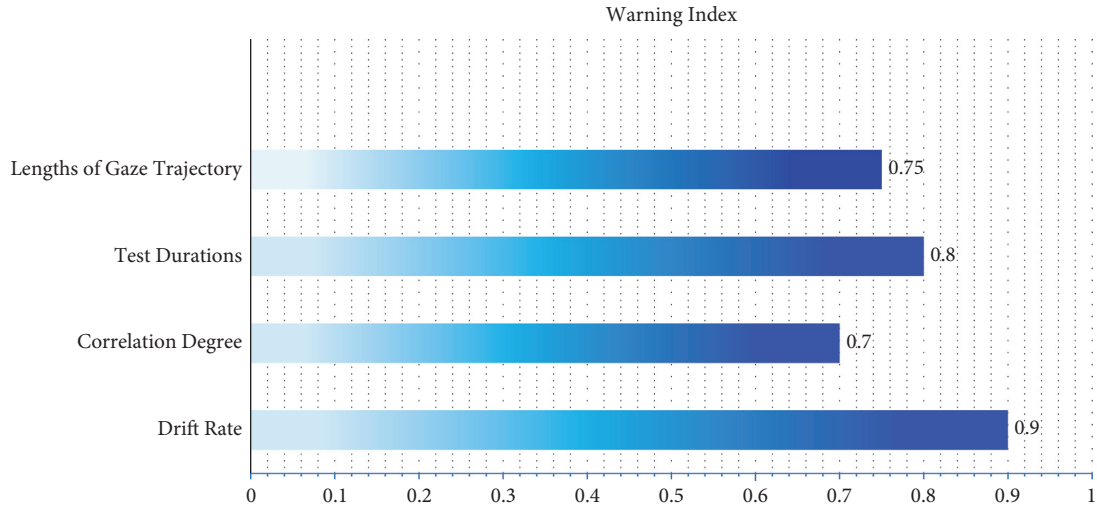


FIGURE 11: Indexes and coefficient with high frequency base on semistructured interview.

TABLE 4: Participant characteristics between MCI group and normal cognitive (NC) group.

Items	Classification	Total population	MCI group	NC group	<i>p</i> value
Gender ( <i>n</i> , %)	Male	6 (18.75%)	2 (33.33%)	4 (66.67%)	0.373
	Female	26 (81.25%)	14 (53.85%)	12 (46.15%)	
Age (years)		84.50 (7.00)	86.00 (3.00)	83.00 (10.75)	0.064
Educational level (years)		12.00(6.75)	12.00 (6.00)	9.00 (9.75)	0.175

TABLE 5: Comparison of three indexes between MCI group and NC group.

Group	Test durations		Lengths of gaze trajectory		Correlation degree		Drift rate	
	<i>M</i> (quartile)	<i>p</i> value	<i>M</i> (quartile)	<i>p</i> value	Mean $\pm$ SD	<i>p</i> value	<i>M</i> (quartile)	<i>p</i> value
MCI group	50.44 (85.89)	0.407	11562.63 (9539.73)	0.274	0.74 $\pm$ 0.19	0.01	0.28 (0.27)	0.019
NC group	43.39 (42.22)		10474.51 (6638.62)		0.87 $\pm$ 0.08		0.16 (0.18)	

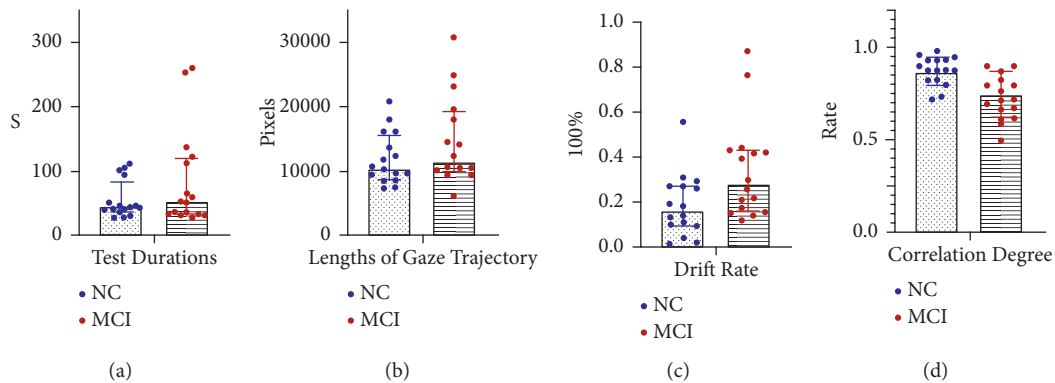


FIGURE 12: Comparison of indexes between MCI and NC. (a) Distribution of test durations. (b) Distribution of lengths of gaze trajectory. (c) Distribution of drift rate. (d) Distribution of correlation degree.

In conclusion, the key technology of human-computer interactive 2.5-minute fast digital early warning for MCI proposed in this paper (1) realizes the collection of eye movement objective data through eye tracker and achieves the extraction and analysis of eye movement data by using the objective assessment key technology, which overcomes the lack of objective assessment data of screening scales in

common use. (2) And by assessing visuospatial function which is an early sensitivity assessment dimension for MCI, it makes up for the defects of low sensitivity and screening rate of screening scales for MCI. Besides, (3) based on the dynamic automatic paradigm scene combined with human-computer interaction, it overcomes the shortcomings of long time consumption and requiring professionals to assess

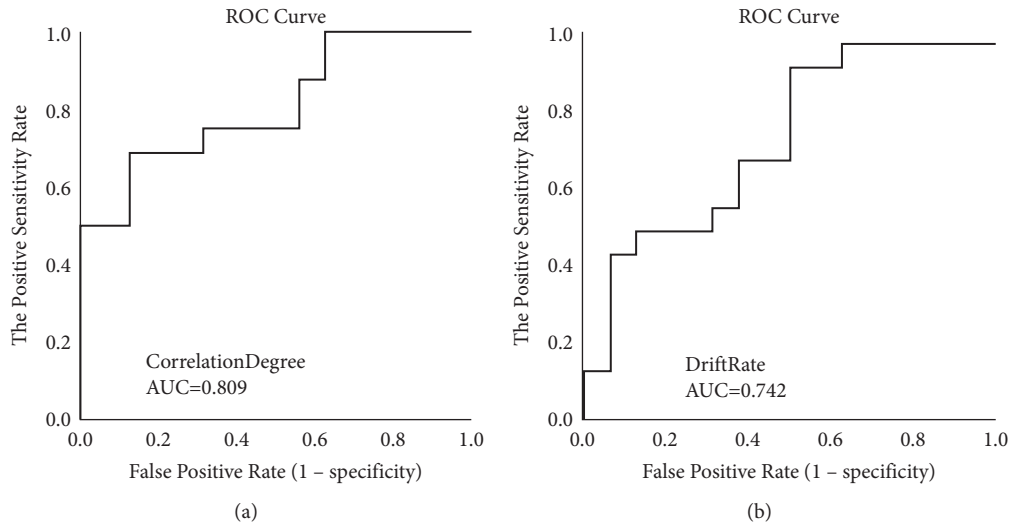


FIGURE 13: ROC curve of single index for MCI early warning. (a) ROC of correlation degree. (b) ROC of drift rate.

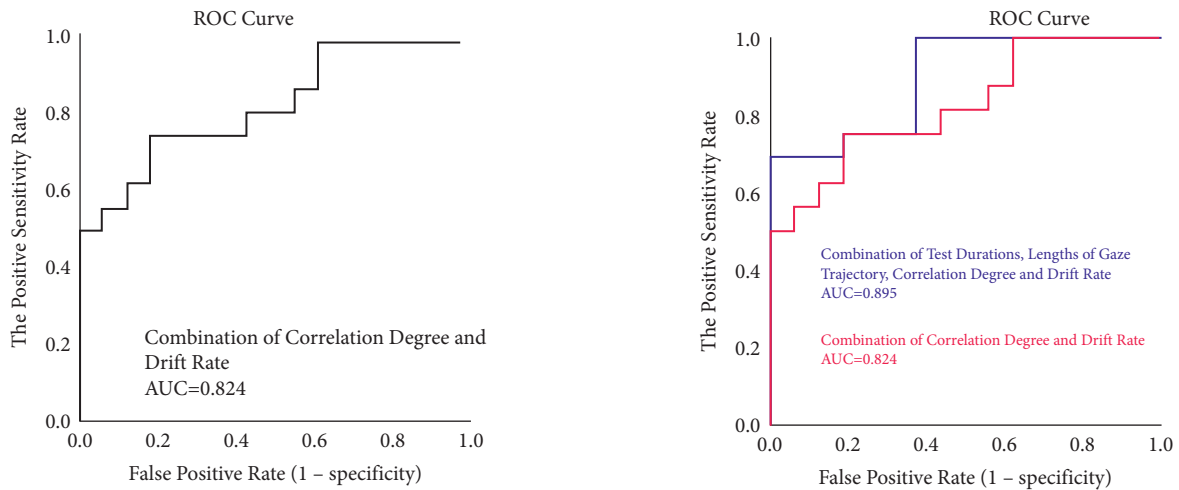


FIGURE 14: ROC curve of joint indexes for MCI early warning.

FIGURE 15: ROC curve of combing two indexes with discriminative significance and two key indexes for MCI early warning.

screening scales and previously developed computer systems, which makes it easier for subjects to obtain an activated test experience in the test. Early warning in the elderly population through experimental verification realizes the fast and high-precise MCI screening and early warning in 2.5 minutes among the elderly population and constructs a new model for MCI early warning based on “correlation degree” and “drift rate.”

There are three shortcomings of this paper. (1) The subjects of this study are in the age range of 70–90 years, who are generally older. It is found that the sensitivity is low when carrying out the human-computer interaction paradigm for some people under the age of 60. Therefore, in the later stage, we will develop grouping devices of different ages to automatically adjust the early warning sensitivity according to different ages. (2) In this paper, the subjects only cover MCI and normal cognitive aged, not including AD population. (3) The total sample size of this experiment is limited, and further experiments need to be carried out to expand the

universality of the research results. Future research will explore the utility of this technology in the mild cognitive impairment, and AD population of different ages and establish free MCI human-computer interactive early warning program and public case database for all researchers in <https://brainhealth.zcmu.edu.cn/en/mciw>.

### Data Availability

The datasets used in the experiments and discussed in the paper are available from the corresponding author on reasonable request.

### Ethical Approval

The present study was approved by Medical Ethics Committee of Zhejiang Chinese Medical University (Approval no. 20210806-1).

## Consent

All participants gave written informed consent, and identifiable participants gave written consent to publish their images.

## Conflicts of Interest

None of the authors have conflicts of interest related to the research and results presented in this paper.

## Authors' Contributions

Nan Li and Xiaotong Yang These authors contributed equally to this work and should be considered co-first authors

## Acknowledgments

This research is supported by “Integrated equipment package for individualized intelligent rehabilitation diagnosis and treatment of traditional Chinese medical encephalopathy at home-Take insomnia and epilepsy as an example based on traditional Chinese medical intelligent syndrome differentiation for treatment” which is the Modern TCM Diagnosis and Treatment Equipment R&D Project (2021C03116) in the Key R&D Program of Zhejiang Province and “A Full Cycle Intelligent Rehabilitation System for Future TCM Multimodal Perception Enhanced Diagnosis and Integrated Treatment of Medicine and Medical Equipment for Alzheimer’s Disease” which is a scientific research project (2021JKZDZC04) of Zhejiang Chinese Medical University.

## References

- [1] S. Gauthier, B. Reisberg, M. Zaudig et al., “Mild cognitive impairment,” *The lancet*, vol. 367, no. 9518, pp. 1262–1270, 2006.
- [2] P. S. Sachdev, D. M. Lipnicki, N. A. Kochan et al., “The prevalence of mild cognitive impairment in diverse geographical and ethnocultural regions: the COSMIC collaboration,” *PLoS One*, vol. 10, no. 11, Article ID e0142388, 2015.
- [3] Alzheimer’s Association, “2016 Alzheimer’s disease facts and figures,” *Alzheimer’s and Dementia*, vol. 12, no. 4, pp. 459–509, 2016.
- [4] C. A. Lindbergh, R. K. Dishman, and L. S. Miller, “Functional disability in mild cognitive impairment: a systematic review and meta-analysis,” *Neuropsychology Review*, vol. 26, no. 2, pp. 129–159, 2016.
- [5] M. S. Albert, S. T. DeKosky, D. Dickson et al., “The diagnosis of mild cognitive impairment due to Alzheimer’s disease: recommendations from the National Institute on Aging-Alzheimer’s Association workgroups on diagnostic guidelines for Alzheimer’s disease,” *Alzheimer’s and Dementia*, vol. 7, no. 3, pp. 270–279, 2011.
- [6] T. I. M. Standish, D. W. Molloy, A. Cunje, and D. L. Lewis, “Do the ABCS 135 short cognitive screen and its subtests discriminate between normal cognition, mild cognitive impairment and dementia?” *International Journal of Geriatric Psychiatry*, vol. 22, no. 3, pp. 189–194, 2007.
- [7] R. F. Buckley, K. P. Sparks, K. V. Papp et al., “Computerized cognitive testing for use in clinical trials: a comparison of the NIH Toolbox and Cogstate C3 batteries,” *The journal of prevention of Alzheimer’s disease*, vol. 4, no. 1, pp. 3–11, 2017.
- [8] Y. C. Youn, J. M. Pyun, N. Ryu et al., “Use of the clock drawing test and the rey-osterrieth complex figure test-copy with convolutional neural networks to predict cognitive impairment,” *Alzheimer’s Research & Therapy*, vol. 13, no. 1, pp. 1–7, 2021.
- [9] S. C. Seligman and T. Giovannetti, “The potential utility of eye movements in the detection and characterization of everyday functional difficulties in mild cognitive impairment,” *Neuropsychology Review*, vol. 25, no. 2, pp. 199–215, 2015.
- [10] K. K. Alichniewicz, F. Brunner, H. H. Klünemann, and M. W. Greenlee, “Neural correlates of saccadic inhibition in healthy elderly and patients with amnesic mild cognitive impairment,” *Frontiers in Psychology*, vol. 4, p. 467, 2013.
- [11] A. Peltsch, A. Hemraj, A. Garcia, and D. P. Munoz, “Saccade deficits in amnesic mild cognitive impairment resemble mild Alzheimer’s disease,” *European Journal of Neuroscience*, vol. 39, no. 11, pp. 2000–2013, 2014.
- [12] H. W. Heuer, J. B. Mirsky, E. L. Kong et al., “Antisaccade task reflects cortical involvement in mild cognitive impairment,” *Neurology*, vol. 81, no. 14, pp. 1235–1243, 2013.
- [13] S. C. Beh, B. Muthusamy, P. Calabresi et al., “Hiding in plain sight: a closer look at posterior cortical atrophy,” *Practical Neurology*, vol. 15, no. 1, pp. 5–13, 2015.
- [14] J. Gu, G.-H. Li, X.-W. Liu, Q. Xu, and H.-M. Zhu, “Diagnostic value of exploratory eye trajectory movement in mild cognitive impairment and early Alzheimer’s disease,” *Journal of Jiangsu University: Medical Edition*, vol. 27, no. 02, pp. 146–149, 2017.
- [15] M. Phillips, E. Boman, H. Österman, D. Willhite, and M. Laska, “Olfactory and visuospatial learning and memory performance in two strains of Alzheimer’s disease model mice-A longitudinal study,” *PLoS One*, vol. 6, no. 5, Article ID e19567, 2011.
- [16] Y. Cheng and S.-F. Xiao, “Research progress of mild cognitive impairment in China (English),” *Shanghai psychiatry*, vol. 26, no. 01, pp. 4–14, 2014.
- [17] M. M. Carrasquillo, Q. u. A. Khan, M. E. Murray et al., “Late-onset Alzheimer disease genetic variants in posterior cortical atrophy and posterior AD,” *Neurology*, vol. 82, no. 16, pp. 1455–1462, 2014.
- [18] F. W. Bylsma, D. X. Rasmusson, G. W. Rebok, P. M. Keyl, L. Tune, and J. Brandt, “Changes in visual fixation and saccadic eye movements in Alzheimer’s disease,” *International Journal of Psychophysiology*, vol. 19, no. 1, pp. 33–40, 1995.
- [19] R. U. Haque, C. M. Manzanares, L. N. Brown et al., “VisMET: a passive, efficient, and sensitive assessment of visuospatial memory in healthy aging, mild cognitive impairment, and Alzheimer’s disease,” *Learning & Memory*, vol. 26, no. 3, pp. 93–100, 2019.
- [20] E. Bartoli, F. Caso, G. Magnani, and G. Baud-Bovy, “Low-cost robotic assessment of visuo-motor deficits in Alzheimer’s disease,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 7, pp. 852–860, 2017.
- [21] I. E. Plattner, L. Mbakile-Mahlanza, S. Marobela et al., “Developing a computerized brief cognitive screening battery for Botswana: a feasibility study,” *Archives of Clinical Neuropsychology*, vol. 34, no. 5, pp. 682–689, 2019.
- [22] K. V. Papp, A. Samaroo, H. C. Chou et al., “Unsupervised mobile cognitive testing for use in preclinical Alzheimer’s

- disease,” *Alzheimer’s and Dementia: Diagnosis, Assessment & Disease Monitoring*, vol. 13, no. 1, Article ID e, 2021.
- [23] P. Maruff, Y. Y. Lim, D. Darby et al., “Clinical utility of the cogstate brief battery in identifying cognitive impairment in mild cognitive impairment and Alzheimer’s disease,” *BMC psychology*, vol. 1, no. 1, pp. 30–11, 2013.
- [24] A. Oyama, S. Takeda, Y. Ito et al., “Novel method for rapid assessment of cognitive impairment using high-performance eye-tracking technology,” *Scientific Reports*, vol. 9, no. 1, pp. 1–9, 2019.
- [25] L. E. Hebert, J. Weuve, P. A. Scherr, and D. A. Evans, “Alzheimer disease in the United States (2010–2050) estimated using the 2010 census,” *Neurology*, vol. 80, no. 19, pp. 1778–1783, 2013.
- [26] G.-D. Batty, Q. Li, R. Huxley et al., “Oral disease in relation to future risk of dementia and cognitive decline: prospective cohort study based on the Action in Diabetes and Vascular Disease: Preterax and Diamicron Modified-Release Controlled Evaluation (ADVANCE) trial,” *European Psychiatry*, vol. 28, no. 1, pp. 49–52, 2013.
- [27] K. K. F. Tsoi, J. Y. C. Chan, H. W. Hirai, S. Y. S. Wong, and T. C. Y. Kwok, “Cognitive tests to detect dementia,” *JAMA Internal Medicine*, vol. 175, no. 9, pp. 1450–1458, 2015.
- [28] J. Kim, H. K. Na, J. Byun et al., “Tracking cognitive decline in amnesic mild cognitive impairment and early-stage Alzheimer dementia: mini-mental state examination versus neuropsychological battery,” *Dementia and Geriatric Cognitive Disorders*, vol. 44, no. 1-2, pp. 105–117, 2017.
- [29] J. Elvesjo, M. Skogo, and G. Elvers, “Method and installation for detecting and following an eye and the gaze direction thereof,” U.S. Patent and Trademark Office, No. 7,572,008, Washington, DC, USA, 2009.
- [30] S. Sayago, A. Rosales, V. Righi, S. M. Ferreira, G. W. Coleman, and J. Blat, “On the conceptualization, design, and evaluation of appealing, meaningful, and playable digital games for older people,” *Games and Culture*, vol. 11, no. 1-2, pp. 53–80, 2016.
- [31] A.-H. Guo and X.-P. Pan, “Eye tracking study of Alzheimer’s disease,” *Guangdong medicine*, vol. 42, no. 9, pp. 1132–1135, 2021.

## Retraction

# Retracted: RFID Data Analysis and Evaluation Based on Big Data and Data Clustering

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] L. Lv, "RFID Data Analysis and Evaluation Based on Big Data and Data Clustering," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3432688, 10 pages, 2022.

## Research Article

# RFID Data Analysis and Evaluation Based on Big Data and Data Clustering

Lihua Lv 

*School of Information and Technology, Zhejiang Institute of Economics and Trade, Hangzhou, Zhejiang, China*

Correspondence should be addressed to Lihua Lv; [littledrop@zjiet.edu.cn](mailto:littledrop@zjiet.edu.cn)

Received 30 January 2022; Accepted 2 March 2022; Published 26 March 2022

Academic Editor: Daqing Gong

Copyright © 2022 Lihua Lv. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The era people live in is the era of big data, and massive data carry a large amount of information. This study aims to analyze RFID data based on big data and clustering algorithms. In this study, a RFID data extraction technology based on joint Kalman filter fusion is proposed. In the system, the proposed data extraction technology can effectively read RFID tags. The data are recorded, and the KM-KL clustering algorithm is proposed for RFID data, which combines the advantages of the K-means algorithm. The improved KM-KL clustering algorithm can effectively analyze and evaluate RFID data. The experimental results of this study prove that the recognition error rate of the RFID data extraction technology based on the joint Kalman filter fusion is only 2.7%. The improved KM-KL clustering algorithm also has better performance than the traditional algorithm.

## 1. Introduction

Since human society entered the information age, information technology-related industries have rapidly developed, driving changes in many fields such as computers, communications, and data storage. With the advent of this change, the main content of people's daily work has gradually changed to the processing of various data information. With the support of software and hardware technology, a huge amount of data in various forms has become an important part of people's lives. However, these massive data information often have no obvious structural characteristics at first glance, and traditional methods and ideas are difficult to achieve the purpose of fully understanding and understanding them. Coupled with the huge amount of data and rich types of data, traditional statistical methods cannot fully extract effective information, and it becomes more difficult to obtain valuable information in the data. On top of this predicament, data mining technology has received extensive attention. With the help of data mining technology, people can extract originally hidden, unknown, and valuable information from a large amount of disorganized data. By using it in other fields, the value of information has been fully reflected. Data mining technology is a multidisciplinary field technology, which has rapidly

developed under the huge demand for information data processing technology in the past ten years.

This study mainly analyzes the reasonable solution to the problem of time and space requirements and tries to propose an effective clustering scheme for RFID data. Compared with previous studies, this study mainly has the following two innovations: (1) based on big data technology, the RFID data extraction technology has been improved, and the RFID data extraction technology based on joint Kalman filter fusion is proposed. (2) An improved KM-KL clustering algorithm is proposed.

*1.1. Related Work.* RFID technology is one of the important technologies of the internet of things and big data. It realizes the transmission between things, and a large amount of data is generated during the transmission process. There are many research studies at home and abroad. Wang and Jiang proposed an OCT prediction method utilizing a combination of order and real-time shop floor RFID data. It uses accurate RFID data to describe the real-time load situation of the job shop and tries to mine the mapping relationship between RFID data and OCT from historical data [1]. Fazzinga et al. investigated a method for interpreting RFID



data in the context of object tracking. It consists of converting readings generated by RFID-tracked moving objects into semantic locations on a map by exploiting some integrity constraints [2]. Martinus et al. created a system for supermarket shopping research that enables people to scan items themselves and quickly pay. They focus on using RFID technology in the system [3]. Microcontrolled devices were studied by Novikov et al., and their aim was to develop a personal portable effective dose dosimeter with an RFID data channel without built-in power supply [4]. In the process of analysis, cluster analysis is often used for massive RFID data. Yunoh et al. focused on the analysis of fatigue strain signals based on clustering and classification methods. They grouped the feature extraction using the K-means clustering method to obtain the appropriate number of datasets. The classification process is performed by using an artificial neural network (ANN) for optimal pattern recognition. Experiments show that their algorithm is about 92% accurate [5]. Wang et al. predict natural disasters by modeling meteorological disasters. They clustered natural disasters through the detention analysis method, summarized the characteristics of each weather system, and designed disaster control projects based on this. Their experiments show that control engineering has a high centrality with the occurrence of disasters, and engineering implementation can reduce disasters [6]. Chen used MATLAB software to perform statistical analysis and cluster analysis on the daily PM<sub>(2.5)</sub> concentrations observed in Shanghai in 2014. The results show that the PM<sub>(2.5)</sub> concentration in spring and winter is higher than that in summer and autumn, and the annual distribution of PM<sub>(2.5)</sub> concentration is U-shaped [7]. Balik et al. compared health indicators and health expenditures in 28 European Union (EU) countries, 6 EU candidate countries, and 3 European Free Trade Association (EFTA) countries using a cluster analysis method. As a result of the cluster analysis, the countries were divided into 3 clusters, the first cluster including Turkey had the lowest average per capita public, private, and out-of-pocket health expenditure of the three clusters [8]. It can be easily seen from the related research that the research on RFID technology is more on its application level than on data analysis, so the cluster analysis for RFID data is rarely studied.

## 2. RFID Technology

Wireless identification technology, also known as radio-frequency identification technology, is often referred to as RFID (radio-frequency identification) technology. This communication technology originated from the identification application of British fighter jets during World War II and has been commercialized since the 1960s. In particular, the promotion and application of the US Food and Drug Administration (FDA) and Walmart supermarkets, such as real-time detection of fresh commodities in supermarkets, intelligent price tag system, inventory control system, and intelligent shopping cart, have greatly expanded the application market of RFID technology in the world [9, 10].

The RFID system is generally composed of three parts: reader (Reader), electronic label (Tag), and application software. The tag can be divided into two parts: an antenna and a special chip, and the chip is attached with a unique identification code, indicating the basic information attached to the object. The principle is that the reader transmits the emitted radio-frequency signal to the electronic tag by means of electromagnetic or inductive coupling, so as to drive the electronic tag circuit to transmit its internal data to the reader. The reader accepts and interprets the relevant data in sequence and uses the software system for relevant processing [11]. Radio-frequency identification has the characteristics of noncontact, not affected by environmental factors, a large amount of stored information, readable and writable, fast recognition speed, long recognition distance, and anticollision function. It can process multiple radio-frequency cards at the same time, and the two identification methods of RFID are shown in Figure 1:

The peculiar way in which RFID devices acquire data leads to serious uncertainty in their data [12, 13]. There are three main reasons for the uncertainty of RFID data, namely, missing reading, overreading, and dirty data. Compared with the three, multireading and dirty data phenomena are more contingent and less likely to occur. The phenomenon of missing reading is relatively common, which is the main factor leading to the uncertainty of RFID data.

*2.1. The Principle of Antong Data Collection.* Because RFID technology has the advantages of fast and real time, it also has extensive analysis in the field of its data collection. The RFID data acquisition system generally includes three parts: the target with the electrical label, the RFID base station, and the information center. The system composition is shown in Figure 2:

Based on the working principle of RFID technology, the RFID data acquisition system performs radio-frequency identification on the target equipped with electronic tags and transmits the identification data to the traffic information center. Its working principle is as follows: (1) the system enters the acquisition state from the dormant state and (2) determines whether to stop the acquisition by judging the stop flag bit of reading the radio-frequency tag. If there is no stop command, the system starts to read the target information. (3) The RFID base station judges whether a tag enters the radio-frequency field through radio-frequency technology. (4) After sending the request, the mark is selected to confirm whether to detect the specific target. (5) After the verification is passed, the resident writer starts to read the tag carried by the target and returns and transmits the read response data to the information center after the operation is successful. The workflow of the RFID data acquisition system is shown in Figure 3:

*2.2. Joint Kalman Filter Fusion of RFID Technology and Coil Technology.* For the fusion methods of multisensor information of the same detection section, the fusion methods suitable for dynamic traffic parameters mainly include joint

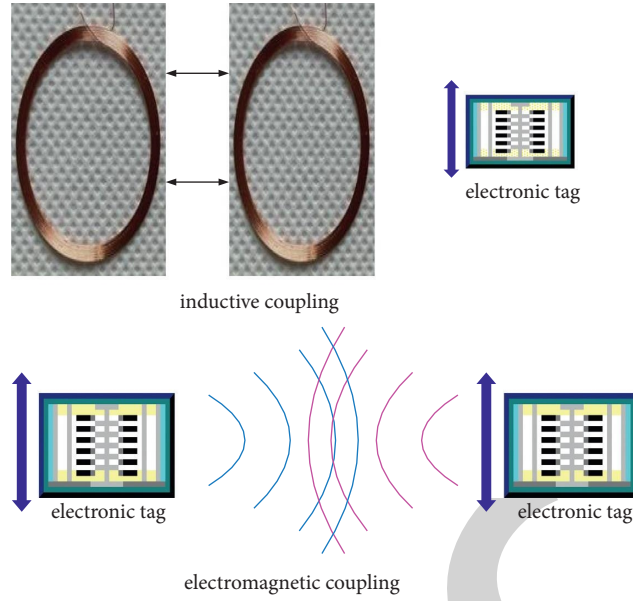


FIGURE 1: RFID (radio-frequency identification) technology.

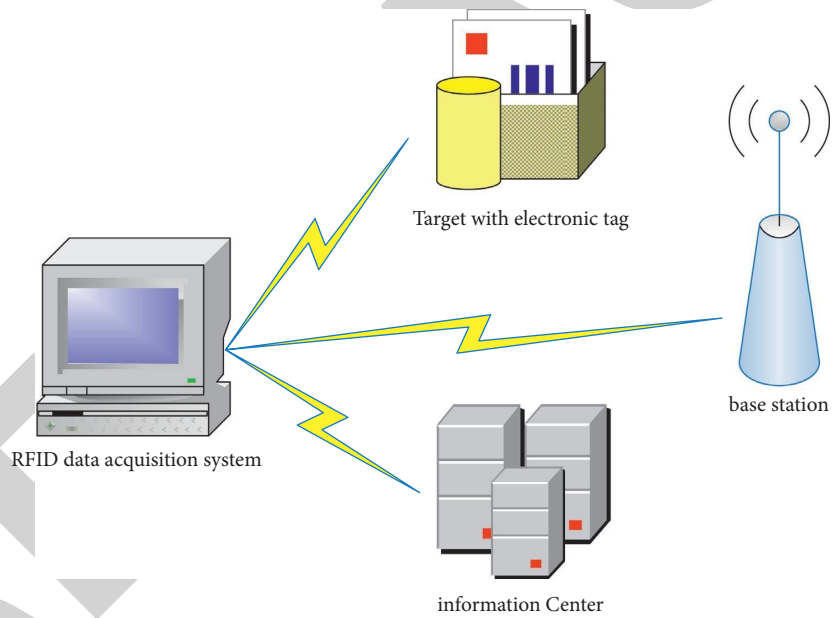


FIGURE 2: Composition of RFID traffic data collection system.

filter fusion, centralized filter fusion, and neural network-based fusion methods.

The basic structure of the joint Kalman filter adopts the method of two-stage processing of data and scattered filtering. The combined filter consists of the main filter and several subfilters.

The structure diagram of the joint Kalman filtering algorithm used in this study is shown in Figure 4:

In the structure of the joint Kalman filter in Figure 4, the subfilter first performs independent filtering, transmits the filtering result to the main filter, and at the same time completes the optimal fusion of the collected information.

Among them, the state formula and measurement formula of the filter are equations (1) and (2), respectively:

$$X_i(k) = \Phi(k, k-1) \cdot X_i(k-1) + W_i(k-1), \quad (1)$$

$$Y_i(k) = H_i(k) \cdot X_i(k) + V_i(k). \quad (2)$$

where  $W_i(k-1)$  is the noise of the dynamic model, and its covariance is  $Q_i(k-1)$ .  $V_i(k)$  is the observation noise, and its covariance is  $R_i(k)$ . Since the traffic volume collected by the coil sensor is slightly smaller than the actual value, this study improves the joint Kalman filter. In this study, a sensor with a known acquisition error is selected as the reference

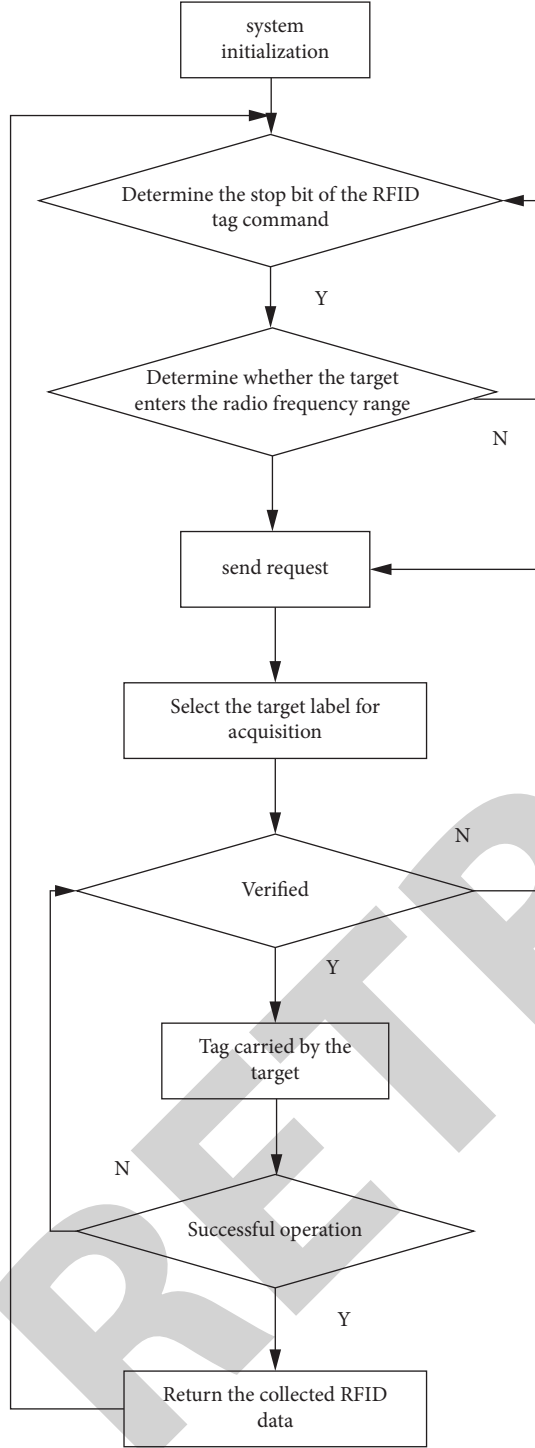


FIGURE 3: RFID traffic data collection system operation process.

sensor, and the method of the first comparison and then fusion is adopted. Its filtering formula is shown in equations (3) and (4):

$$X_i(k+1) = \Phi(k+1, k) \cdot X_i(k), \quad (3)$$

$$P_i(k+1, k) = \Phi(k+1, k)P_i(k)\Phi^T(k+1, k) + Q_i(k). \quad (4)$$

The main filter does not perform filtering but directly performs data fusion. The fusion method is shown in equations (5) and (6):

$$X_g = P_g \left( \sum_{i=1}^{n,m} P_i^{-1} X_i \right), \quad (5)$$

$$P_g = \left( \sum_{i=1}^{n,m} P_i^{-1} \right)^{-1}; \quad i = 1, 2. \quad (6)$$

When the main filter is fused, it feeds back information to the subfilters according to the fusion result, where the feedback factors are shown in equations (7) and (8):

$$\beta_i = (p_1^{-1} + p_2^{-1})^{-1} \cdot p_i^{-1}, \quad (7)$$

$$p_i^{-1}(k) = \beta_i \cdot p_g^{-1}(k). \quad (8)$$

When the condition of  $Q1 < Q2$  is not satisfied, it can be seen from equations (5) to (7) that  $\beta_1 + \beta_2 = 1$ .  $X_g = \beta_1 X_1 + \beta_2 X_2$ , that is, the fusion result  $X_g$  will take a value between  $X_1$  and  $X_2$ , and change between the two values as the feedback factor changes. At the same time, the subfilters are fed back according to the fusion result, and the information is redistributed to improve the fusion accuracy.

Compared with the centralized Kalman filter and artificial neural network, the joint Kalman filter has the advantages of flexible design, a simple algorithm, better fault tolerance, and more suitable for real-time systems.

In this study, the data  $Q1$  collected by RFID technology and the data  $Q2$  collected by the coil sensor are directly combined with Kalman filtering (hereinafter referred to as general fusion). At the same time, it compares the size of  $Q1$  and  $Q2$  and then fuses them according to the fusion method (hereinafter referred to as improved fusion), as shown in Figure 5.

As can be seen from Figure 5, when the traffic volume collected by the RFID technology and the coil sensor is smaller than the actual value, the fusion effect is poor. When the traffic volume collected by RFID technology is smaller than that of the coil sensor, the traffic volume collected by the coil sensor is directly taken as the fusion value. The fusion curve coincides with the traffic volume change curve collected by the coil sensor, which significantly improves the accuracy of the fusion result.

It can be seen from Table 1 that in dataset 1, the error of the improved fusion is the same as that of the general fusion, so the improved fusion method in this study has certain applicable conditions. That is, when the traffic volume collected by RFID technology is smaller than the traffic volume collected by the known small sensor or larger than the traffic volume collected by the known large sensor, the improved fusion proposed in this study produces better results. From dataset 2 and dataset 3, it can be seen that the relative error is significantly reduced by using the improved method in this study for fusion.

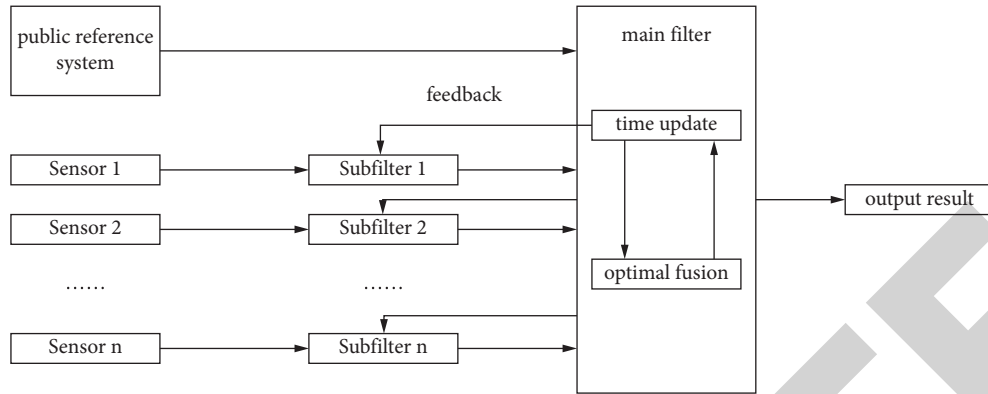


FIGURE 4: Structure diagram of joint Kalman filter.

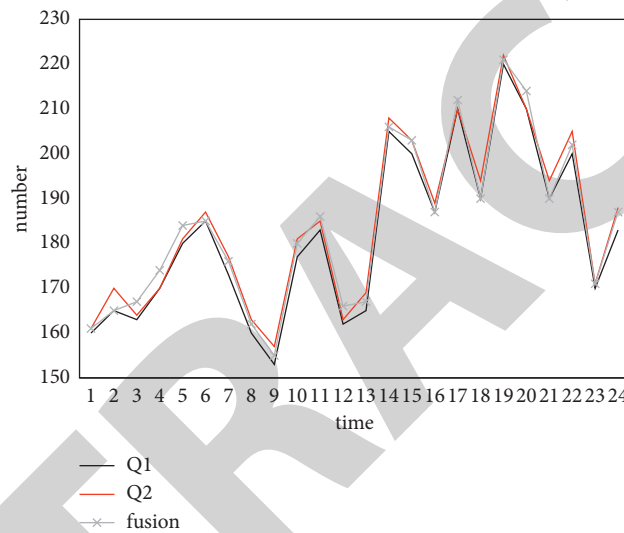


FIGURE 5: Traffic fusion results of the joint Kalman filter algorithm.

TABLE 1: Comparison of average relative errors (unit: %).

Data	Traditional RFID technology	Coil sensor	General fusion	Improve fusion
1	8.52	6.27	0.97	0.97
2	10.04	4.93	7.23	4.6
3	9.56	5.78	5.72	2.74
Average	9.37	5.66	4.64	2.77

### 3. Analysis and Result

3.1. Cluster Analysis. Pattern recognition is to study the automatic processing and interpretation of patterns through the use of mathematical techniques by computers, and the environment and objects are collectively referred to as “patterns.” With the development of computer technology, it is possible for human beings to study the complex information processing process. An important form of the process is the recognition of the environment and objects by the living body. In the field of pattern recognition and statistical analysis, cluster analysis has always been the focus and research direction of the academic community. So far, a large number of theories and methods have been proposed, and remarkable

research results have been achieved. After a long period of research and development, clustering analysis can be divided as follows [14, 15]:

3.1.1. Partition-Based Approach. The typical algorithm is the K-means algorithm and its characteristics are summarized in Table 1, so it will not be repeated here.

3.1.2. Hierarchical Approach. Hierarchical clustering methods use top-down splitting or bottom-up agglomeration to represent datasets in a hierarchical tree structure. Each object is first treated as a separate cluster, and these clusters are merged into larger and larger clusters until a

given end condition is met. This clustering method is a bottom-up agglomerative clustering method. Divisive hierarchical clustering is just the opposite, where all objects are first grouped into one large cluster. Then, it is gradually divided into smaller and smaller clusters according to the similarity, until the end condition is satisfied.

**3.1.3. Density-Based Methods.** Density-based methods are divided into different clusters according to different aggregation densities in data object sets, and clusters with similar densities are divided into one cluster. However, its calculation is large, usually is  $O(n^2)$ . In addition, based on the density, the setting of parameters has a great influence on the performance of the algorithm, and there is no good solution at present, which mainly depends on the user's experience to select parameters.

**3.1.4. Grid-Based Approach.** Grid-based clustering method divides the data space into a certain number of ultra-rectangular grid cells according to the partition parameters, maps the data objects to the corresponding grid cells, and then merges the adjacent grid cells into a connected region, that is, a cluster. Grid-based clustering method has good scalability for the size of datasets, can handle large-scale datasets, and can find clusters with arbitrary shapes. Grid-based clustering method is usually combined with the density-based clustering method [16].

**3.1.5. Model-Based Approach.** The model-based clustering method is to use a specific model for cluster analysis and try to optimize the fit between the actual data and the model. Neural network-based clustering methods and statistical learning-based clustering methods are two types of model-based clustering methods. Among the neural network models commonly used in cluster analysis are the self-organizing map (SOM) model [17], the adaptive resonance theory model [18], and the learning vector quantization (LVQ) model. The EM clustering algorithm based on the Gaussian mixture model is a typical clustering method based on statistical learning [19].

**3.1.6. Fuzzy Clustering Method.** The clustering method described above can be regarded as hard clustering. Different from hard clustering, fuzzy clustering is a soft clustering method. Fuzzy C-means clustering algorithm [20] is a popular fuzzy clustering algorithm at present, which transforms the clustering problem into an optimization problem and uses an iterative method to solve it. It has a simple design, good clustering performance, and wide application. However, the algorithm is sensitive to the initial conditions, easily falls into local optimum, requires a large amount of computation, and has a low resolution for objects in the overlapping area of class boundaries.

Table 2 summarizes the clustering methods commonly used at present, and the advantages and disadvantages of each method, in which NCA means to determine the number of classes sensitive to initial values and outliers in advance.

### 3.2. Partition-Based Clustering Algorithm—KM-KL Algorithm

**3.2.1. The Basic Concept of KM-KL Algorithm.** In the traditional partition-based clustering algorithm, the set of uncertain objects is divided into given  $K$  clusters according to their mutual distances. The KM-KL algorithm in this study follows the idea of the traditional partition-based clustering algorithm and divides the uncertain dataset  $O$  containing  $n$  uncertain objects into  $k$  clusters. They are denoted as  $C_{1,L}, C_k$ , respectively, use the symbol  $C_i$  as the center point of each cluster, and use the center point to represent the cluster  $C_i$ . In the partition-based clustering algorithm, the following properties are included as follows:

- (1) All clusters belong to the uncertain dataset, that is  $C_i \subseteq O (1 \leq i \leq k)$ .
- (2) Each cluster contains at least one data record, that is  $C_i \neq \varphi$ .
- (3) Each data record belongs to one cluster and belongs to only one cluster, that is, when equation (9) is satisfied, equation (10) can be satisfied:

$$U_{i=1}^k C_i = O, \quad i \neq j, \quad (9)$$

$$C_i \cap C_j = \varphi. \quad (10)$$

On the basis of traditional partition-based clustering, the KM-KL algorithm in this study uses KL divergence as a similarity measure.

The algorithm divides uncertain objects into  $k$  clusters and selects an optimal cluster center point for each cluster to minimize the sum of the overall KL divergence. In the clustering algorithm, in the division formed by a clustering process, the sum of the KL divergence among all objects is shown in the following:

$$\text{TKL} = \sum_{i=1}^k \sum_{P \in C_i} D(PC_i). \quad (11)$$

This formula is used to measure the quality of the clustering. The larger the TKL value, the worse the quality of this clustering is, and the smaller the TKL value is, the better the quality of this clustering is. In equation (11),  $D(PC_i)$  represents the KL distance from the object  $P$  to the cluster center point  $C_i$ , which is used to assign the object to each cluster.

$$\sum_{P \in C_i} D(PC_i). \quad (12)$$

Formula (12) represents the sum of the KL distances from each object  $P$  in the cluster  $C_i$  to the cluster center point  $C_i$ , according to which the structure of the cluster  $C_i$  can be adjusted.

**3.2.2. KM-KL Algorithm Description.** The KM-KL algorithm in this study is an extension of the K-medoids algorithm using KL divergence as a similarity measure. This algorithm

TABLE 2: Types of clustering methods.

Type	Advantage	Insufficient
Based on partition method	Wide application, fast convergence, incremental clustering, and suitability for large-scale data	It is necessary to determine the NCA, which is sensitive to initial values and outliers, so as to find circular clusters
Hierarchy-based method	It does not need to determine the NCA and can find clusters of any shape, which is suitable for data of any attribute and has strong clustering ability	No backtracking, no exchange of data objects between classes, no full processing of large-scale data, and no incremental clustering
Density-based method	It does not need to determine the NCA, can find clusters of different shapes, can detect outliers, and has good adaptability to large datasets	It is very sensitive to parameters. For datasets with uneven density distribution, the quality of clustering results is not high

clusters uncertain data on probabilistic similarity [11–22]. The algorithm is divided into two parts, the initial clustering division stage and the center replacement stage.

(1) *Initializing the clustering stage.* In the initial clustering and division stage, the algorithm selects  $k$  cluster center points one by one to initialize the clustering and division of other objects. The first center point  $C_1$  selects the object with the smallest sum of KL divergence of other objects in the uncertain object set  $O$ , that is, formula:

$$C_1 = \arg \min \left( \sum_{P' \in O} D(P'P) \right), \quad (13)$$

The remaining  $k - 1$  center points are iteratively selected. At the  $i$ -th iteration, the algorithm selects the object  $C_i$  that minimizes TKL. When it is calculated that  $P'$  will be allocated to the cluster with  $P$  as the new cluster center point, the contribution of this allocation to reducing TKL is shown in the following:

$$\max(0, \min_{j=1}^{i-1} (D(P'C_j) - D(P'C_i))) \quad (14)$$

From formula (14), it can be calculated the sum of the reduction in TKL by the division of all unselected objects in this round, which is expressed as  $\text{DEC}(P)$ . In the KM-KL algorithm, the center point selected by the  $i$ -th iteration is the object with the largest  $\text{DEC}(P)$ , that is, formula:

$$C_i = \arg \max_{P \in O \{C_1, \dots, C_{i-1}\}} (\text{DEC}(P)). \quad (15)$$

This ensures that the initial cluster center point selected at each time is optimal, because this selection can minimize the dissimilarity between clusters. The initial clustering division phase ends when all  $k$  cluster center points are selected and then enters the second phase of the algorithm, the center replacement phase.

(2) *Center replacement stage.* In the center replacement stage, the algorithm iteratively replaces the cluster center points with all noncenter points in the cluster and selects the optimal cluster center point to improve the clustering quality. Every time redistribution occurs, it records the TKL reduction value after redistribution and selects the object with the largest value as the new cluster center, as shown in the following:

$$P_{\max} = \arg \max_{P \in O \{C_1, \dots, C_{i-1}\}} (\text{DEC}(P)). \quad (16)$$

At this time, if it is  $\text{DEC}(P_{\max})$ , it means that this round of exchange improves the clustering effect. Otherwise, the algorithm ends and the final clustering is generated. The example shown in Figure 6 is used to illustrate the KM-KL algorithm flow.

As shown in Figure 6, Figure 6(a) shows the distribution of the uncertain object set. Supposing that the uncertain objects  $a_1$ ,  $a_2$ , and  $a_3$  satisfy the same distribution, it forms a cluster. The uncertain objects  $b_1$ ,  $b_2$ , and  $b_3$  satisfy another distribution and it forms another cluster, which is the clustering result of the benchmark. Figure 6(b) shows the KL divergence between uncertain objects and the initialized TKL value. Due to the asymmetry of KL divergence, in Figure 6(a), for example, the divergence from  $a_1$  to  $a_2$  is not equal to the divergence from  $a_2$  to  $a_1$ . At this time, the KM-KL algorithm is used to cluster the uncertain objects in Figure 6(a), and the input parameter  $k$  is equal to 2.

Finally, the KM-KL algorithm forms two clusters  $\{a_1, a_2, a_3\}$  and  $\{b_2, b_1, b_3\}$  with  $a_1$  and  $b_2$  as the center points, respectively, and the division is the optimal division.

*3.3. Algorithm Performance Test and Effectiveness Analysis.* The experimental platform is configured as follows: Intel(R) processor, 2.94 GHz main frequency, 2G memory, using Windows7 operating system. The experimental program is written in C++ language, compiled and run on VS2010, and simulated by MATLAB.

This study generates datasets on discrete and continuous domains, respectively. In the continuous domain, an uncertain object is a sample taken from a continuous distribution, and in the discrete domain, the dataset is transformed from a continuous model. The continuous domain is discretized using a meshing method, dividing each dimension into two equal parts, thus dividing the  $d$ -dimensional data space into the same 2D cells. The probability of an object in a certain unit is the sum of the probabilities of all sample points of the object in this unit. The value range of the data is  $[0, 1]^d$ . Three different probability distributions will be used in this study, namely, uniform distribution, Gaussian distribution, and inverse Gaussian distribution, which is generated by Gaussian distribution, as shown in Figure 7.



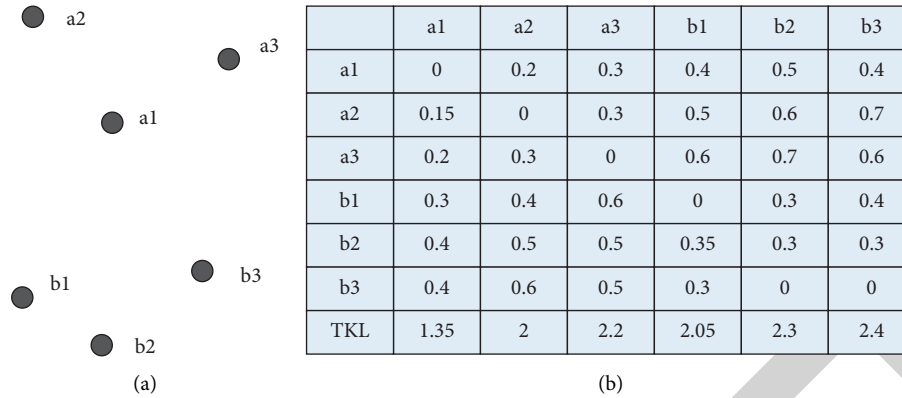


FIGURE 6: Data objects and their KL divergence.

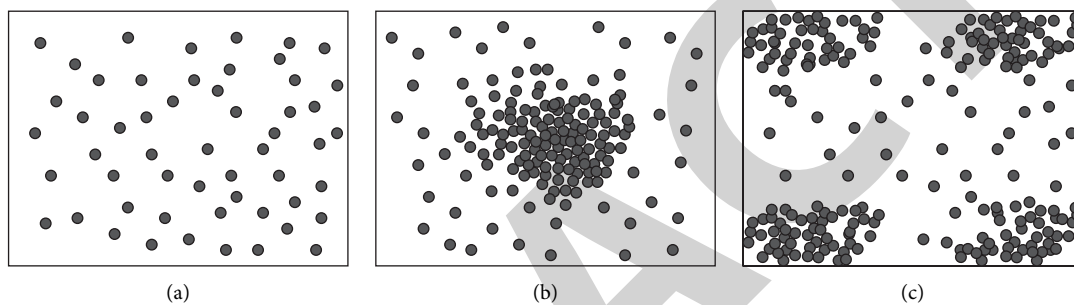


FIGURE 7: Three different distributions. (a) Evenly distribution. (b) Normal distribution. (c) Inverse Gaussian distribution.

The experiments in this study are mainly divided into three parts: first, the effect of the clustering algorithm using KL divergence is analyzed. Second, it shows the improvement of the computational efficiency of the clustering algorithm in this study after using an efficient implementation method and an improved fast Gaussian transform. Finally, experiments show the scalability of the algorithm on large datasets.

In classical partition-based and density-based clustering algorithms, KL divergence and geometric distance are used as similarity measures, respectively, to compare clustering quality. Among the partition-based clustering algorithms, the UK-means algorithm using geometric distance (denoted as UK) and the KM-KL and RKM-KL algorithms proposed in this study are compared. Among the density-based clustering algorithms, the FDBSCAN algorithm using geometric distance (denoted as FD) and the DB-KL algorithm in this study are compared.

In the experiment, the base number of objects is set to  $n = 100$  by default, each object contains  $s = 100$  sample points, the data dimension  $d = 4$ , and the default setting of clustering produces  $k = 6$  clusters. For the density-based algorithm,  $k$  is not used as a parameter, and the parameters' density threshold  $\mu$  and distance radius  $\varepsilon$  need to be set here. Among them, the density threshold is  $\mu = 5$  according to the recommendation, and the distance radius  $\varepsilon$  is continuously adjusted in the experiment so that the z FD and DB-KL can produce clustering results' approximate to  $k$  clusters.

Since the complexity of the density-based algorithm DB-KL exponentially increases with respect to the number of

data objects, the algorithm is not suitable for a large number of data objects. The scalability of the RKM-KL algorithm is mainly tested here. The dataset defaults to 4 dimensions and contains 10 clusters.

Figures 8 and 9 show the effect of the RKM-KL algorithm and the RKM-KL-FGT algorithm on the clustering quality of the algorithm when the data object cardinality is large and the number of object samples is large. It can be seen that when the object cardinality is large or the number of samples is large, the algorithm has a similar quality trend as when the amount of data is small. Moreover, as in the previous experimental analysis, the clustering quality of the RKM-KL-FGT algorithm will be reduced to a certain extent compared to the RKM-KL algorithm.

Figure 10 shows the running time of the RKM-KL algorithm and the RKM-KL-FGT algorithm when the dataset is large. It can be seen that the running time of the RKM-KL algorithm linearly increases as the cardinality of the data objects increases. The RKM-KL-FGT algorithm uses an improved fast Gaussian transform to obtain approximations, so the running time of the algorithm is shorter, and the running time smoothly increases with the increase in cardinality. However, since the calculation of the sum of KL-divergence squarely increases with respect to the number of samples of the object, the calculation time of the algorithm RKM-KL rapidly increases with the increase in the sample data volume. While the computation time of the RKM-KL-FGT algorithm is almost unaffected (it grows linearly), obviously, RKM-KL-FGT is more effective when the amount of data is large.

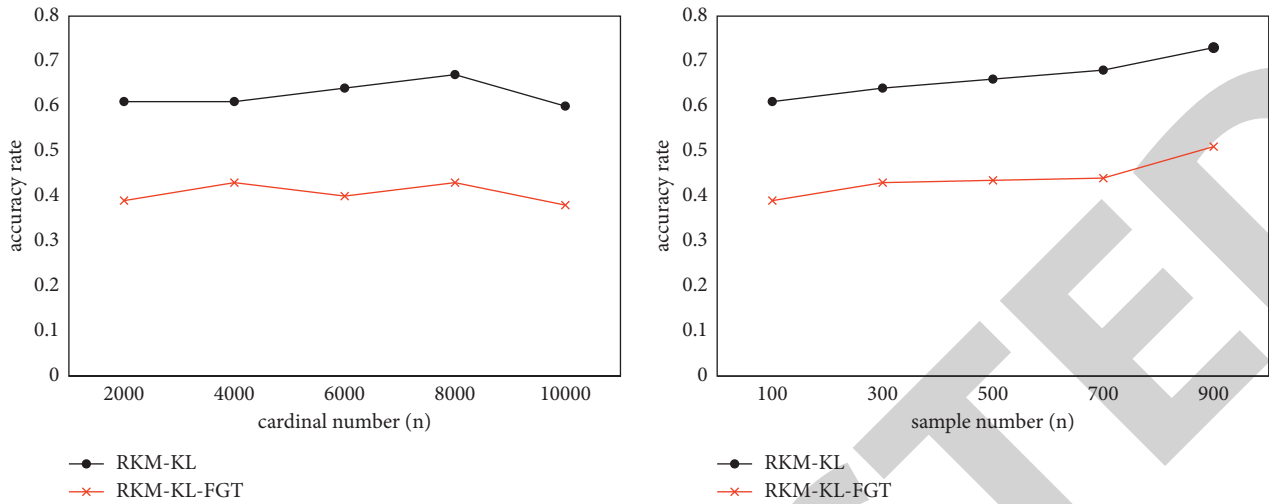


FIGURE 8: Accuracy when using large datasets on continuous domains.

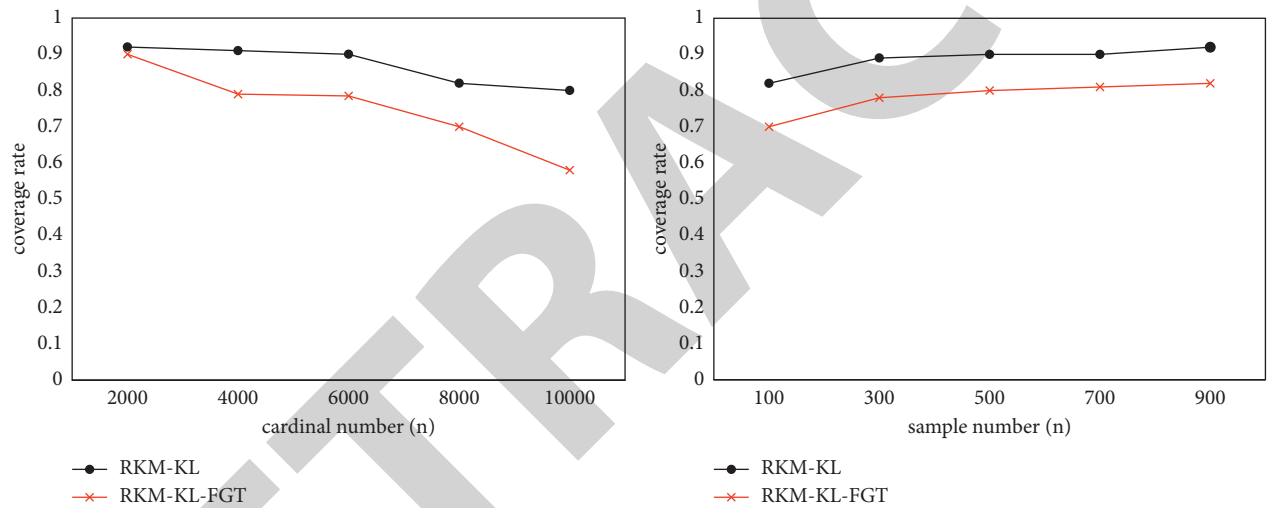


FIGURE 9: Coverage when using large datasets on continuous domains.

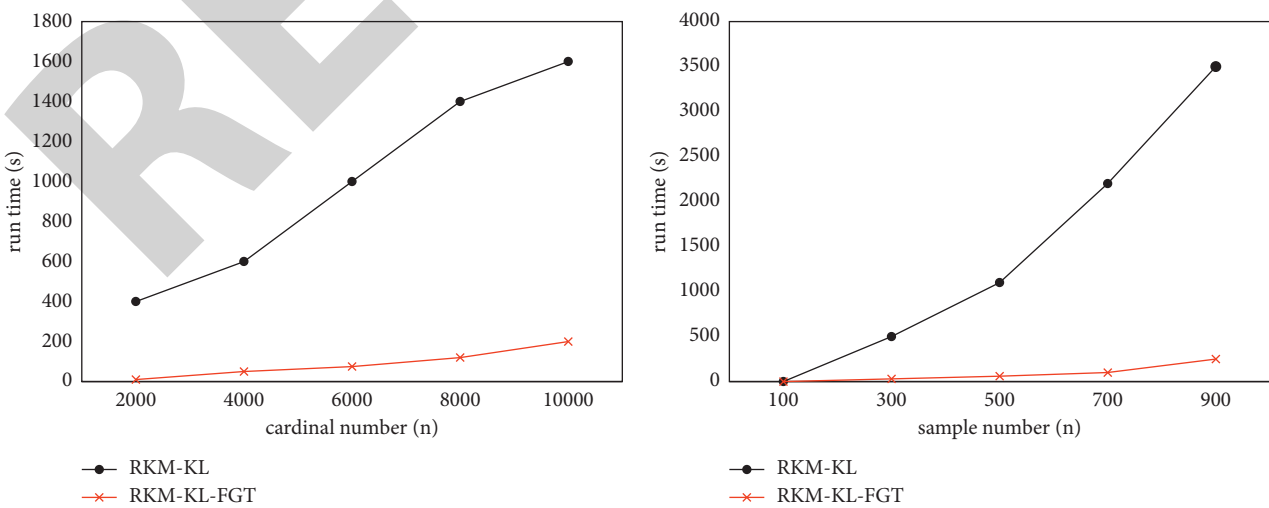


FIGURE 10: Running time when using large datasets on continuous domains.

## *Retraction*

# **Retracted: An Optimized Decision Method for Smart Teaching Effect Based on Cloud Computing and Deep Learning**

### **Computational Intelligence and Neuroscience**

Received 3 October 2023; Accepted 3 October 2023; Published 4 October 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### **References**

- [1] M. Jiang and Y. Sun, "An Optimized Decision Method for Smart Teaching Effect Based on Cloud Computing and Deep Learning," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 6907172, 10 pages, 2022.

## Research Article

# An Optimized Decision Method for Smart Teaching Effect Based on Cloud Computing and Deep Learning

Miaomiao Jiang  and Yuwei Sun 

*Physical Education College of Harbin University of Commerce, Harbin 150028, China*

Correspondence should be addressed to Yuwei Sun; 101885@hrbcu.edu.cn

Received 31 January 2022; Accepted 21 February 2022; Published 26 March 2022

Academic Editor: Daqing Gong

Copyright © 2022 Miaomiao Jiang and Yuwei Sun. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to improve the effect of intelligent teaching and give full play to the role of intelligent technology in modern physical education, in this paper, cloud computing and deep learning methods are used to comprehensively evaluate the teaching effect of colleges and universities, and calculate the evaluation effect and accuracy. Cloud computing and deep learning algorithm combine the teaching evaluation scale, teaching content, and characteristics to formulate teaching plans for different students and realize targeted teaching evaluation. The results show that the teaching evaluation method proposed in this paper can improve students' learning interest by about 30%, enhance learning initiative by about 20%, and the matching rate between the actual teaching effect and the expected requirements is 98%. Therefore, cloud computing and deep learning model can improve the accuracy of teaching effect evaluation in colleges and universities, provide support for the formulation of teaching evaluation schemes, and promote the development of intelligent teaching in colleges and universities.

## 1. Introduction

Cloud computing and deep learning was first applied in the field of power and gradually developed to the field of teaching in the later stage. Cloud computing and deep learning can have a key impact on the improvement of students' learning status and teaching content, and improve the overall teaching effect. In 2017, China Education Commission put forward the strategy of "comprehensive quality education" and the concept of "double reduction," which reduced the college tennis teaching time from 48 hours in 2012 to 32 hours in 2020, with a decrease of 30% [1]. The results are shown in Figure 1. In this paper, tennis intelligent teaching is taken as an example to analyze the advanced nature and intelligent advantages of intelligent teaching. At the same time, the increase of multimedia, online teaching, and practical courses also greatly reduces the teaching hours of college tennis physical education. With the continuous reduction of teaching hours, the teaching pressure of tennis teachers is increasing and even affects the effect of classroom teaching. Therefore, "teaching optimization of cloud computing and deep learning" has been

put on the agenda [2] and has become the research focus of tennis teaching in major universities. Cloud computing is the mainstream algorithm at present. Through the intelligent search of the data in the network, the horizontal teaching effect evaluation can be realized. Deep learning is the in-depth study of teaching contents and methods to realize the vertical analysis of teaching effect. The integration of cloud computing and deep learning can realize the comprehensive analysis of intelligent teaching and improve the accuracy of effect evaluation. Although cloud computing and deep learning methods are traditional teaching evaluation methods, this paper integrates the two and constructs a new theoretical model, which belongs to a new model attempt. The advantage of the first mock exam is very obvious. It can not only expand the scope of the evaluation of the intelligent teaching effect but also increase the depth of assessment and effectively remedy the deficiency of the single model. Foreign scholars began to study cloud computing and deep learning in 2010, and gradually introduced it into the evaluation of intelligent teaching effect. At the same time, foreign countries have achieved good research results in data preprocessing in cloud

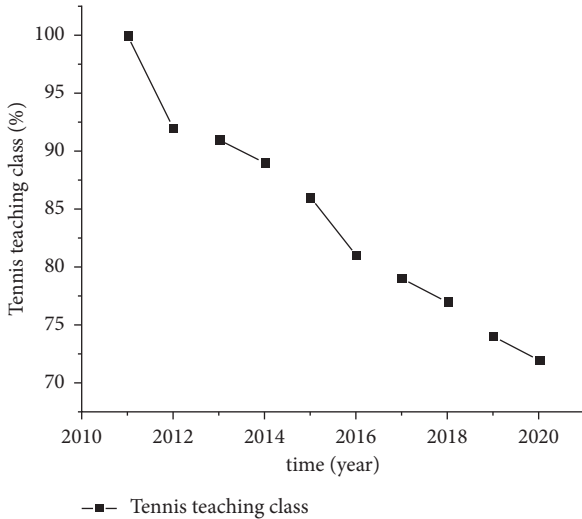


FIGURE 1: Changes in tennis teaching hours.

computing and data mining in deep learning. However, there is a gap between China and foreign countries in terms of intelligent devices and development procedures. Moreover, the national conditions and policies of foreign countries are different from those at home, so it is impossible to copy them. Therefore, domestic scholars should strengthen the research of cloud computing and deep learning, and integrate it with intelligent teaching.

At present, tennis teaching mainly adopts the form of “unified” teaching, which has the advantages of unified teaching content and simple teaching methods, but it cannot meet the “personalized” requirements of students [3]. Under the background of reducing teaching time, increasing teaching pressure, and improving the requirements of quality education, there will be problems of reduced interest in learning and lack of learning law. On this basis, this paper proposes a cloud computing and deep learning model to integrate students’ learning data, classify students’ interests, and analyze the learning results, in order to improve the effect of tennis teaching.

As can be seen from Figures 1 and 2, in college physical education, the demand for tennis courses increases year by year, while the actual course supply decreases year by year. There is a balance between demand and supply in 2015, and the gap will increase from 2015 to 2020. Therefore, tennis teaching has great potential and needs to be solved more accurately.

## 2. Literature Review

There are many research studies on the cloud computing and deep learning model in tennis teaching, but it mainly stays at the theoretical level and lacks practical case studies [4]. Some scholars believed that the increase of tennis content, the application of information technology, and the reduction of teaching time affected the effect of tennis teaching. Some scholars also believed that the intelligent storage student management scale can realize the management of tennis teaching content and can improve the teaching effect of tennis [5]. Some scholars believed that students’ classification and

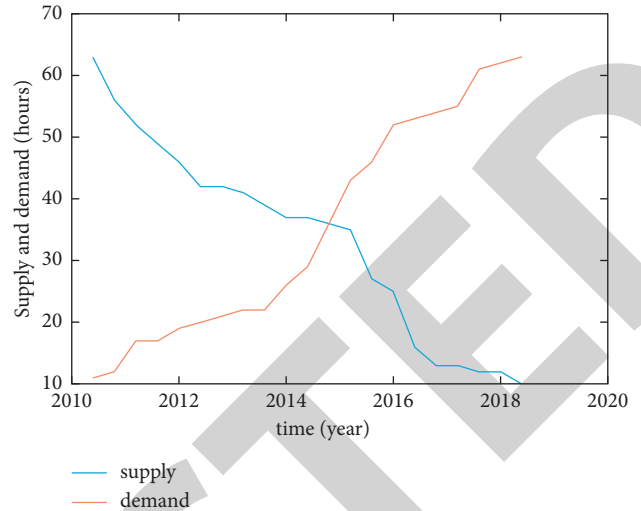


FIGURE 2: The supply and demand of tennis courses.

achievement integration can improve students’ potential and realize personalized teaching [6]. Some scholars believed that the integration of teaching theory and practice and the division of grades can improve the effect of classroom teaching [7]. Some scholars also believed that the intelligent score management scale, score grade, and potential development were the indicators for the evaluation of tennis teaching effect [8], which can realize the comprehensive analysis of students and tennis teaching. In conclusion, scholars at home and abroad believed that the student intelligent score management scale and cloud computing and deep learning method can improve the teaching effect, realize the integration of students’ potential and teaching requirements, and meet the requirements of quality education [9]. However, there was a lack of practical case analysis at home and abroad. On this basis, this paper integrates theory and practice to analyze the effect of cloud computing and deep learning [10].

## 3. Research Model

**3.1. Cloud Computing and Deep Learning Theory.** Cloud computing and deep learning can meet the complex needs of students’ learning and is composed of multiple management units. Cloud computing and deep learning can realize the “personalized service” of students and the resource sharing of tennis teaching [11]. First, the tennis teaching scheme is used to divide the “virtual file” for students in the tennis course management center. Then, a management table is built and stored in each student’s client node  $l$  to form a small-scale data search [12]. Finally, any client node  $l$  calculates the management table of adjacent clients (node  $l+1$ , node  $l-1$ ) to find the required teaching effect data and return the key value of the corresponding client, node  $i$ ,  $\xrightarrow{\text{keyvalue}} \text{DiskTracker}_i(\text{ID}_i)IP_k$  datacenter.

**3.2. Basic Theory of Cloud Computing and Deep Learning.** It is based on the multivariate difference equation, and its mathematical expression is as follows:



$$\begin{cases} \text{node}_i = A \cdot \text{DiskTracker}_i + B \cdot \text{ID}_j + C \cdot \text{IP}_k, \\ \text{keyvalue}_j = D \cdot \text{DiskTracker}_i + E \cdot \text{ID}_j + F \cdot \text{IP}_k + P \sin\left(2k\pi\left(f_e + \frac{f_D}{f_s}\right)\right) \times T(k). \end{cases} \quad (1)$$

Among them, A, B, C, and D are the coefficient values of corresponding parameters in tennis teaching;  $f_e$ ,  $f_s$ , and  $f_D$  are the collected data, sent data, and feedback data;  $P$  is the data amplification factor, which aims to improve the data recognition rate of “abnormal students,” and  $T(k)$  is the data actually received by the center. According to the results of  $f_e$ ,  $f_s$ , and  $f_D$ , the level of students’ tennis teaching is judged. If  $f_e = f_s = f_D \neq 0$  students’ tennis teaching is normal, and if  $f_s \neq f_D = 0$  is near the extreme value, it indicates that students’ tennis teaching is abnormal [13].

### 3.3. Cloud Computing and Deep Learning Process in Tennis Teaching.

- (1) Data access to students’ teaching archives. The access process to students’ teaching archive data can be described by a nonlinear Lagrange equation [14], and the mathematical formula is as follows:

$$g = \frac{d\text{DiskTracker}_i}{d\text{node}_i} \quad (2)$$

$$= -v'(\text{DiskTracker}_i) + F_0 + g(\text{node}_i).$$

Among them,  $g(\text{node}_i)$  is the gap between students’ academic achievements;  $F_0$  represents the collection amount of student effect data;  $F_0$  is a nonlinear discrete function, which  $v'(\text{DiskTracker}_i)$  is the sharing of teaching resources and the satisfaction of individual needs.

- (2) Adjustment of the number of student files. The adjustment of the number of student files is a key factor in the cloud computing and deep learning level, which determines the degree of tennis teaching pressure, and its mathematical calculation formula is as follows:

$$E[g(\text{node}_i)g(\text{node}_i + \alpha)] = 6Kg(\text{node}_i - \alpha) + \xi, \quad (3)$$

where  $K$  is the frequency of tennis resource acquisition;  $\xi$  is the error value of tennis learning achievement change;  $E$  is the adjustment coefficient of the total number of students [15].

- (3) Effective data transmission. The effective data transmission  $F_0$  determines the level of tennis teaching, and its mathematical calculation formula is as follows:

$$F_0 = A_0 \cos(2\pi f_0 \text{node}_i). \quad (4)$$

Among them,  $A_0$  is the function amplitude of students’ achievement change, and  $f_0$  is the frequency of teaching, which together determine the amount of

tennis teaching; teaching effect grade = actual assessment result/expected teaching result \* 100%.

- (4) Shared and personalized service satisfaction. Cloud computing and deep learning requires to evaluate students’ tennis learning effect and meet the sharing of public resources and the satisfaction of individual resources [16], which  $v'(\text{DiskTracker}_i)$  is expressed by a function, and its mathematical calculation formula is as follows:

$$v'(\text{DiskTracker}_i) = \frac{m}{m+n} f_g(\text{DiskTracker}_i)^2 + \frac{n}{m+n} f_h(\text{DiskTracker}_i)^4, \quad (5)$$

where  $m$  and  $n$  are the proportion of shared resources and individual needs in the virtual file, respectively;  $f_g$  and  $f_h$  are a resource sharing function and personality satisfaction function, both of which are internal algorithms in the tennis teaching scheme [17].

- (5) Judgment of critical value. By deriving formula (5), its inflection point can be obtained, that is

$$v''(\text{DiskTracker}_i) = 2 \frac{m}{m+n} f'_g(\text{DiskTracker}_i) + 4 \frac{n}{m+n} f'_h(\text{DiskTracker}_i). \quad (6)$$

After substituting formula (4), it shows that the proportion of shared resources and personalized service resources is in a two-way balance, that is  $A_0 < \sqrt{4m^2 + mn/27n(m+n)}$ , and the proportion between tennis teaching requirements and actual tennis teaching is the best. In order to better analyze the effect of students’ tennis teaching, it will be limited to  $[-\sqrt{4m^2 + mn/27n(m+n)}]$ .

- (6) Identification of abnormal students. The amplification factor is  $\lambda$ , which is added in tennis teaching to identify the abnormal changes of academic performance. If there is a significant change in the scores of any students in the process of tennis teaching, it indicates that the relationship between tennis teaching and students’ learning is normal teaching, otherwise, it is students’ abnormal learning, and the corresponding data are amplified to identify abnormal students.

*3.4. Constraints of Cloud Computing and Deep Learning Model.* In the cloud computing and deep learning of tennis teaching, the original tennis teaching management method is too simple, leaving the analysis of shared resources and



personalized service resources to teachers, increasing the work pressure of teachers, and reducing the management level of students Martins et al. [18]. In view of the above problems, this paper introduces to jointly optimize tennis teaching management and transfer the analysis of some students' teaching to students.

Firstly, metropolis constraint, this paper uses metropolis constraints to judge the effectiveness of student information, so as to reduce the impact of invalid information on the results. If the result is "0," it indicates that the data is invalid, eliminated, or the value is enlarged, otherwise, the corresponding calculation is conducted. After the data is constrained by metropolis, a value is assigned to it to make the constraint result "0" or "1," and the standardization process is completed. According to the above analysis, the mathematical conditions are as follows:

$$\begin{cases} \text{keyvalue}_j \neq \text{keyvalue}_{j+1} \neq 0, \\ \text{keyvalue}_j = \text{keyvalue}_{j+1}, \text{keyvalue}_j \neq 0, \end{cases} \quad (7)$$

where  $M(\cdot)$  is the judgment function of metropolis constraint.

Secondly, the matching between actual teaching and expected teaching effect. The matching between the actual teaching and the expected teaching effect is complex and multidimensional. The actual teaching effect of tennis is not only affected by the number of students but also by the requirements of "sharing degree" and "personality degree." Assuming that the matching of "sharing degree" is  $P_u$  and the matching of "personality degree" is  $P_v$ , formula (5) can be optimized as follows:

$$P = \left\{ \prod \left[ \sum_i^{\infty} v'(\text{DiskTracker}_i) + \varsigma \right], \frac{\kappa(P_u - P_{\min})}{P_{\max} - P_{\min}, P_u < P_v}, \frac{\lambda(P_v - P_{\min})}{P_{\max} - P_{\min}, P_v < P_u} \right\}, \quad (8)$$

where  $P_{\max}$  and  $P_{\min}$  are the maximum and minimum bearing matching values of tennis teaching, the adjustment coefficient of personalized matching, and the matching coefficient of sharing degree.

Thirdly, the effective evaluation and student volume of cloud computing and deep learning. In order to reduce the

overload impact of teaching time on tennis teaching, a part of the analysis of teaching effect is analyzed by students, and the management table is introduced to judge the degree of teaching sharing. The original formulas (2) and (3) are optimized into the following formulas:

$$\begin{cases} R = \int_{i,j}^n \sum_l^{\infty} \lim g(\text{node}_l) [-v'(\text{DiskTracker}_i)] + H_i \cdot f(\widehat{\text{DiskTracker}}_i, \widehat{\text{node}}_l), & R = 1, \\ U = \int_{i,j}^n \sum_l^{\infty} \lim g(\text{node}_l) \left\{ E[g(\text{node}_l)] + H_i \cdot f(\widehat{\text{node}}_l, g(\widehat{\text{node}}_l)) \right\}, \end{cases} \quad (9)$$

where  $R$  is the number of visits to teaching resources, and its value is 0, indicating that teaching resources are not accessed, and students can use the management list to search resources among students;  $U$  is the number of user visits. The number of students in the preset time is compared with the maximum standard required by tennis teaching. If the ratio is  $> 9/10$ , it indicates that the effect of network teaching is good, and the research is strengthened on corresponding teaching contents;  $f(\text{node}_l, g(\text{node}_l))$  is the calculation function of students' misunderstanding rate in tennis delivery;  $H_i$  is the adjustment coefficient of misunderstanding. The judgment process is shown in Figure 3.

Fourthly, the construction steps of cloud computing and deep learning optimization model. According to the student learning scale and the data description formula of tennis

teaching, the simulation model is built. The specific steps are as follows:

- (1) In the cloud computing and deep learning scale, the more frequent the intragroup teaching visits between students through the management scale, the less the pressure of actual tennis teaching and the higher the tennis teaching effect. Therefore,  $\text{node}_i \xleftrightarrow{\text{keyvalue}}_{\text{DiskTracker}_i(ID_j)IP_k}$  datacenter is the best critical value required first, and this value is taken as the stable threshold of the whole tennis teaching. The result of the threshold must be in  $[-\sqrt{m^2 + mn/n(m+n)}, \sqrt{4m^2 + mn/27n(m+n)}]$ .
- (2) Identifying abnormal students in tennis teaching. The abnormality of data transmission in cloud

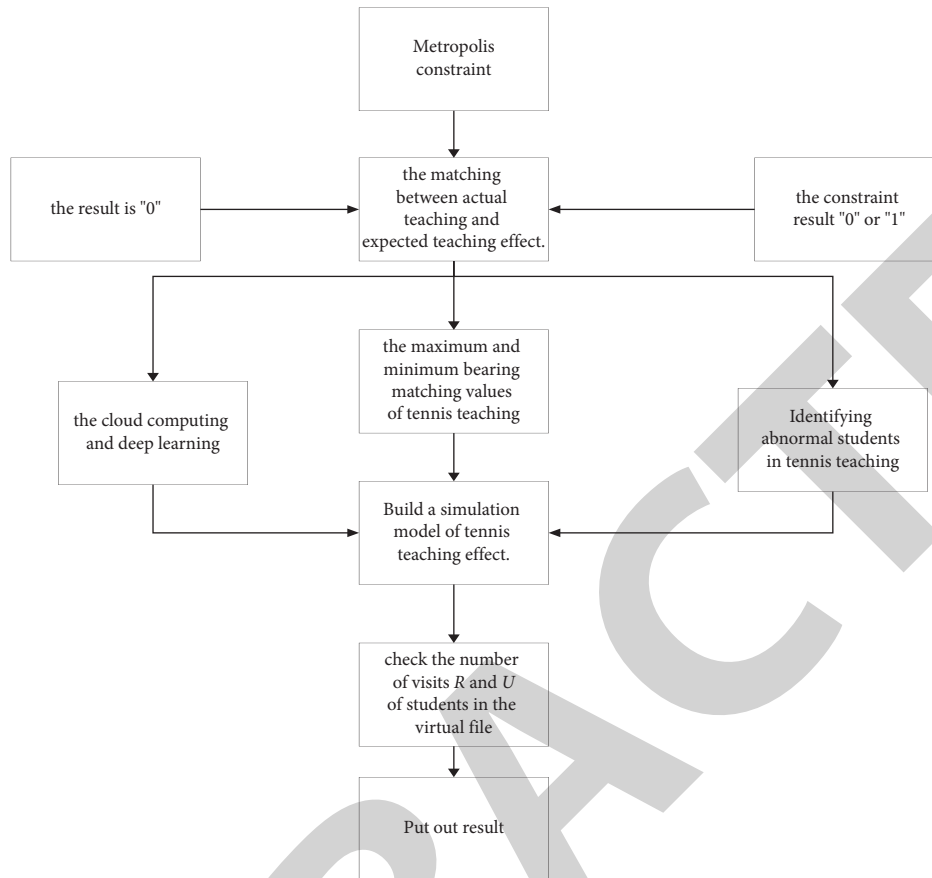


FIGURE 3: The judgment process of tennis courses.

computing and deep learning is an important optimization index. The occurrence of abnormal data is determined by many factors, such as the management table, the integration of DHT and tennis teaching scheme, the division of virtual space in tennis teaching, field type, byte size, access time, and system load. Under the constraint of  $\text{keyvalue}_j \neq \text{keyvalue}_{j+1} \neq 0$ ,  $A_0 < \sqrt{4m^2 + mn/27n(m+n)}$  is the data transmission relatively stable, and the data transmission between tennis teaching and cloud computing and deep learning is normal.

- (3) Building a simulation model of tennis teaching effect. The teaching effect is the final result of tennis teaching and students' understanding, and is affected by the requirements of quality education, teaching time, teaching methods (multimedia, online, and offline), as well as students' own interests. The actual effect of tennis teaching is consistent with the expected requirements, indicating that the effect of tennis teaching meets the requirements. Among them, the calculation of teaching effect is divided into shared resource teaching  $P_u$  and personalized service teaching  $P_v$ . In the calculation process, the number of visits  $R$  and  $U$  of students is checked in the virtual file, and the consistency between the teaching level and the teaching requirements is judged.

#### 4. Case Analysis of Cloud Computing and Deep Learning Model on Tennis Teaching Effect Evaluation

**4.1. Case Introduction.** Taking the 48 h tennis teaching course of a university as an example, 32 classes of offline students are managed. Tennis teaching evaluation needs 1 server, 4 PCs, and 12 mobile phones. The storage mode of students' virtual archives is the combination of disk and solid state, and the storage space is 723 t. The specific circuit diagram is shown in Figure 3. Each student carries out wireless communication through the terminal and sets up their own LAN. Teachers accept students' real-time data and historical data. The data interfaces are filtered signal interface, setting out signal interface, AD conversion interface, and data metering port.

**4.2. Judgment and Comparison of Tennis Teaching Results by Cloud Computing and Deep Learning.** The model built by Simulink in MATLAB software is used for analysis, and the results are shown in Figure 4. Firstly, the data fit. In the cloud computing and deep learning of tennis teaching, the fitting degree between actual teaching and expected teaching is an important management evaluation index. The results are shown in Figure 4.

It can be seen from Figure 4 that the simulation of intelligent teaching is carried out by using a time timer and

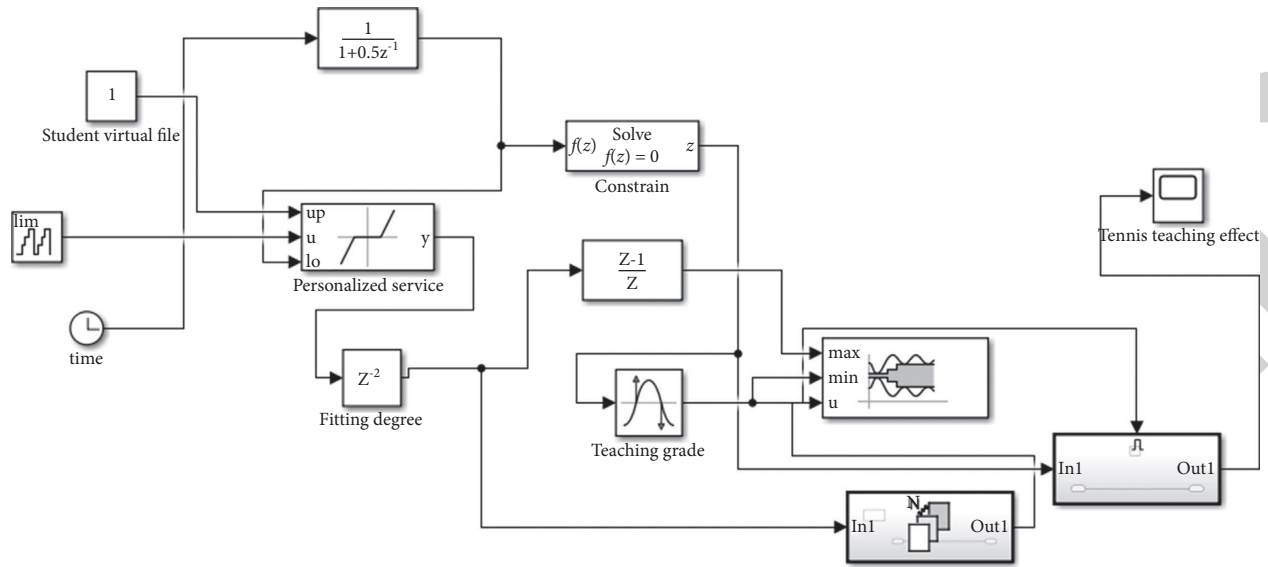


FIGURE 4: The cloud computing and deep learning simulation of college tennis teaching.

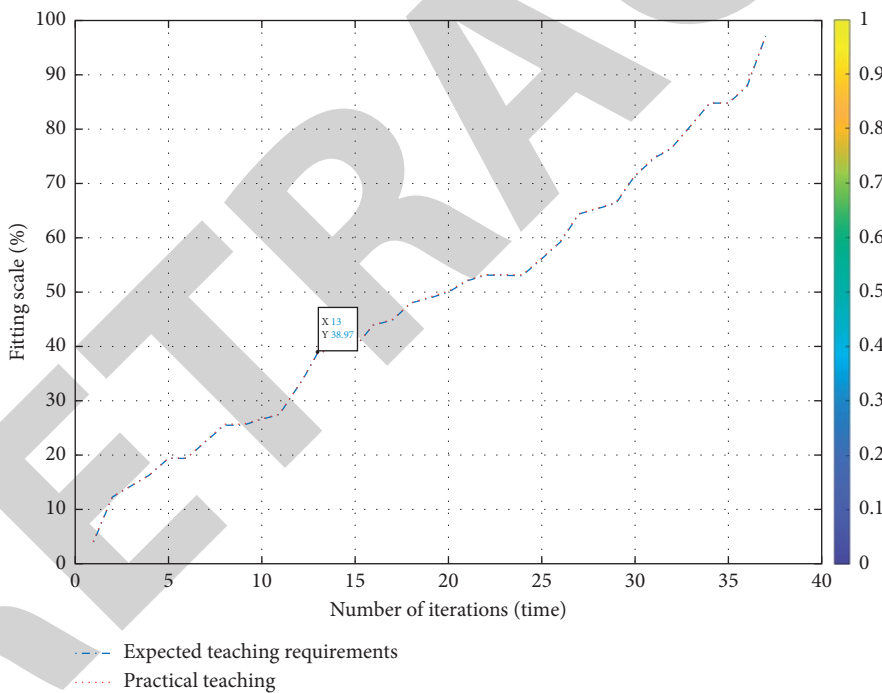


FIGURE 5: The fit ratio of data transmission between actual teaching and expected teaching requirements.

input waveform device; the maximum value, intermediate value, and minimum value are analyzed. Among them, the matching rate between the actual teaching data and the expectation is shown in Figure 5.

It can be seen from Figure 5 that the expected evaluation requirements of intelligent teaching are basically consistent with the actual requirements, indicating that the constructed simulation model can analyze cloud computing and deep learning. In short, the simulation model constructed in this paper is effective and can be used for practical case analysis.

It can be seen from Figures 4 and 5 that the fitting degree  $p$  between the actual teaching and the expected teaching requirements is good. The reason is that before judging, the tennis teaching effects are carried out, and the amplification effects are analyzed. In addition, the teaching data in Figure 6 is relatively stable and the slope changes less. At the same time, the change range of the actual teaching data is relatively stable, which shows that the data processing of the actual teaching is relatively stable,

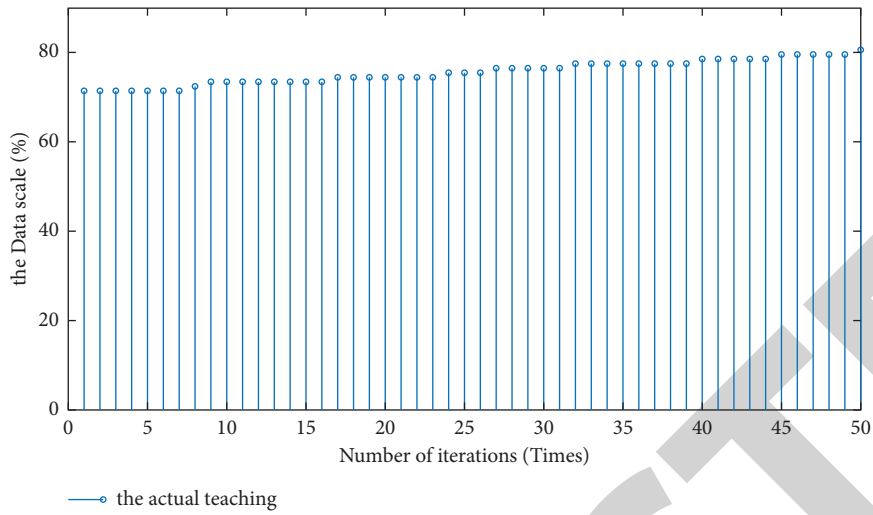


FIGURE 6: The changes in actual teaching data.

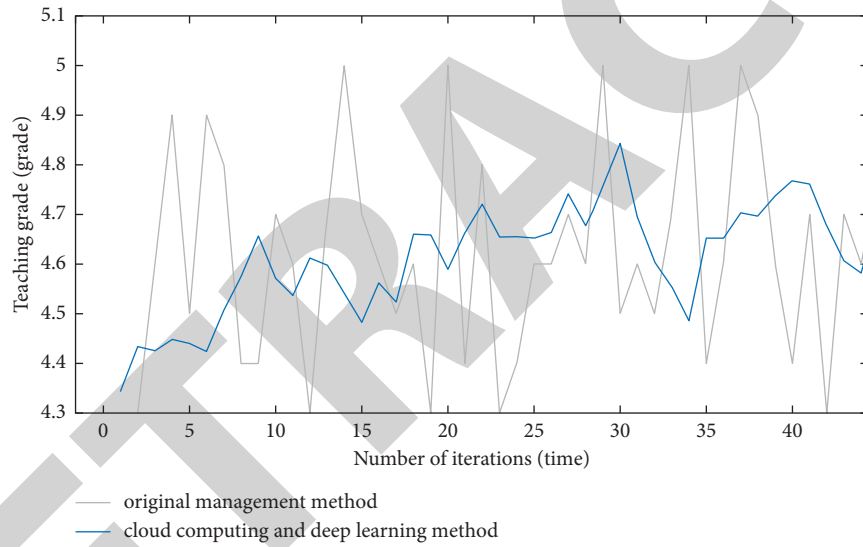


FIGURE 7: The comparison between cloud computing and deep learning and the original management method.

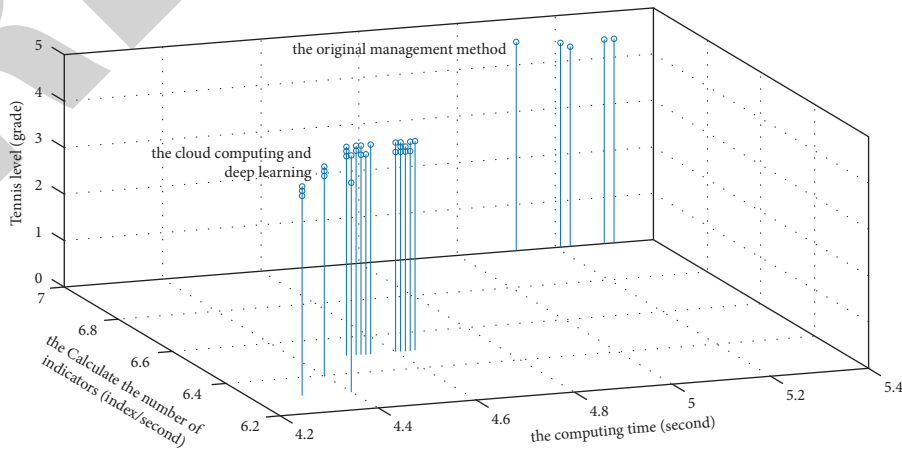


FIGURE 8: The comparison of teaching effect levels of different methods.

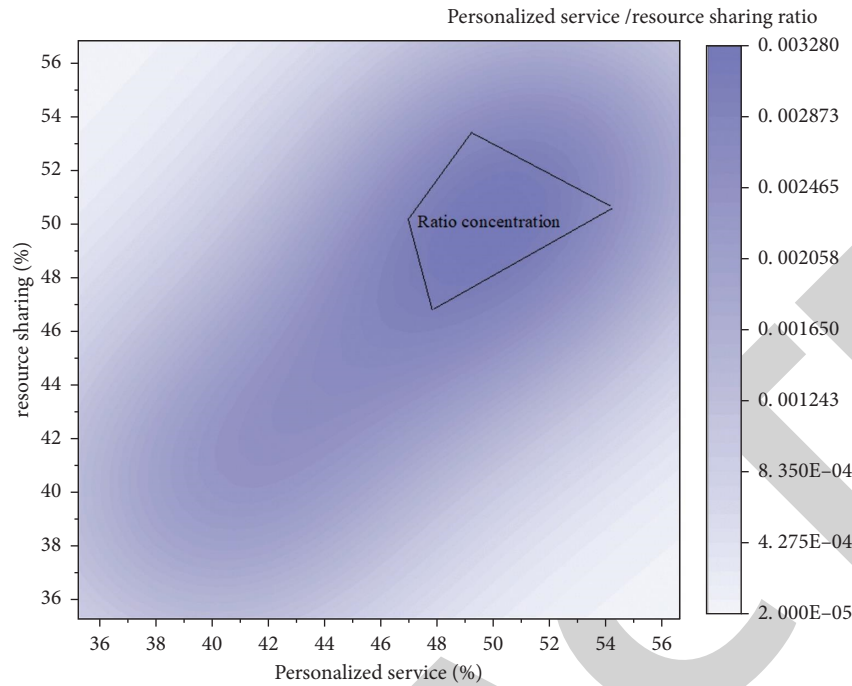


FIGURE 9: The personalized service/resource sharing ratio of the two algorithms.

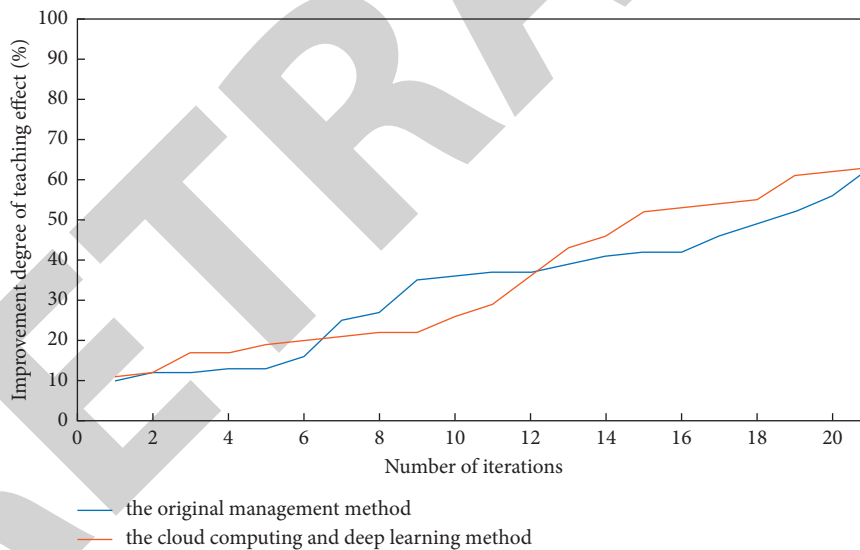


FIGURE 10: The improvement degree of teaching effect of the two algorithms.

which further shows the effectiveness of the algorithm. The reason is that the “group” built by students themselves and the sharing of tennis teaching resources can realize students’ autonomous learning and more efficient understanding of tennis content, and continuously improve their own knowledge combined with offline teaching content.

Secondly, the level of tennis teaching effect. Tennis teaching level is the focus of cloud computing and deep learning, which involves the content of practical operation and theoretical students. Compared with the traditional centralized mode, cloud computing and deep learning mode

has a higher level of teaching effect. On the whole, the evaluation data of actual teaching effect changes little, and the overall comparison process meets the evaluation requirements of learning effect. The reason is due to the self-help learning among students and the effective utilization of teaching resources. The results are shown in Figure 7.

As can be seen from Figure 7, the teaching level under the cloud computing and deep learning method is higher, and the teaching level is also higher than the original evaluation method. At the same time, the fluctuation range of the two algorithms is different. The fluctuation range of

the algorithm proposed in this paper is small and meets the requirements of teaching effect evaluation. The comparison of teaching levels of different methods is shown in Figure 8.

As can be seen from Figure 8, the cloud computing and deep learning method is higher than the original management method in terms of personalized service/resource. Compared with the original method, the algorithm proposed in this paper can evaluate the teaching effect more effectively. The main reason is the horizontal analysis of cloud computing and the vertical analysis of deep learning so as to realize the comprehensive judgment of the effect of intelligent teaching. The comparison of the personalized service/resource sharing ratio of different methods is shown in Figure 9.

As can be seen from Figures 9 and 10 that the calculation results of the two algorithms are higher in terms of service and resource sharing. Compared with the original method, the algorithm proposed in this paper mainly focuses on the core area of the radiograph. Therefore, the personalized strategy and resource sharing of the algorithm proposed in this paper are better than the original algorithm sharing ratio. This shows that under the same teaching time, the cloud computing and deep learning method can provide a higher proportion of personalized services. As we all know, personalized service is one of the important contents of quality education. The cloud computing and deep learning method improves the processing efficiency of tennis teaching on relevant teaching contents and reduces tennis teaching time according to the optimization methods such as the management table so as to better provide personalized management services to students.

## 5. Conclusion

Cloud computing and deep learning can evaluate the effect of tennis teaching in colleges and universities [19]. Cloud computing and deep learning can make up for the deficiency of single teaching and improve the accuracy of tennis teaching effect evaluation. At the same time, cloud computing and deep learning can shorten the evaluation time of teaching effect and make tennis teaching better meet the intelligent requirements. Cloud computing and deep learning is a management method based on middle school teaching. This method comprehensively evaluates the tennis teaching effect by preprocessing the tennis teaching evaluation data so that the evaluation results are more in line with the actual requirements. Cloud computing and deep learning methods reduce the amount of preprocessed data and evaluate the effect faster by identifying "abnormal data." MATLAB simulation results show that under cloud computing and deep learning, the fitting degree  $p$  between the actual tennis teaching effect and the expected teaching effect is better, and the tennis teaching effect level and the ratio of personalized service/resource sharing are higher than the original teaching method. Therefore, in the case of limited teaching time and complex teaching content, the cloud computing and deep learning method can effectively improve the tennis teaching effect and realize the personalized training of students.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported by the Heilongjiang Province Art Science Planning Project (2020B069); the key project of Harbin University of Commerce's 2021 school-level teaching reform and teaching research project (HSDJY202129 (Z)).

## References

- [1] M. M. Alam and F. Faisal, "Public-private partnership (PPP) projection Bangladesh: current status and challenges," *Journal of System and Management Sciences*, vol. 8, no. 4, pp. 38–56, 2021.
- [2] I. Hut, B. Jetic, L. Matija, and Z. Cojbasic, "Machine learning classification of cervical tissue liquid based cytology smear images by optomagnetic imaging spectroscopy," *Tehnicki vjesnik-Technical Gazette*, vol. 26, no. 6, pp. 32–33, 2021.
- [3] A. Burinskiene, "Pharma supply chain: efficiency modelling approach," *Journal of System and Management Sciences*, vol. 8, no. 2, pp. 65–73, 2021.
- [4] Y. Huang, "Analysis of power big data development under the situation of energy Internet," *Modern industrial economy and informatization*, vol. 10, no. 11, pp. 72–73, 2021.
- [5] T. K. Brahmachary, S. Ahmed, and M. S. Mia, "Health,safety and quality management practices in construction sector: a case study," *Journal of System and Management Sciences*, vol. 8, no. 2, pp. 47–64, 2021.
- [6] Y. Sönmez, H. Kutlu, and E. Avci, "A novel approach in analyzing traffic flow by extreme learning machine method," *Tehnicki vjesnik-Technical Gazette*, vol. 26, no. 1, pp. 107–113, 2021.
- [7] P. Marwedel, "Survey on education for cyber-physical systems," *IEEE Design & Test*, vol. 37, no. 6, pp. 56–70, 2020.
- [8] P. Mukucha, "The effects of business process outsourcing on the quality of catering services in tertiary education industry in Zimbabwe," *Cogent Business & Management*, vol. 7, no. 1, pp. 32–34, 2020.
- [9] R. S. Nagoyitsyn, "Interactive technologies in developing student's motivation in physical education and sport," *International Journal of Applied Exercise Physiology*, vol. 9, no. 6, pp. 78–85, 2020.
- [10] T. W. Shen, "Development and evaluation of virtual reality induction electricity prevention education and training tools for construction industry," in *Proceedings of the 7th IEEE International Conference on Consumer Electronics-Taiwan (ICCE-Taiwan)*, pp. 6–11, Taoyuan, Taiwan, September 2020.
- [11] Z. Zhou, X. Zhang, J. Lu, J. Huang, L. Zhou, and Y He, "Energy and economic analysis of life cycle zero energy building in the temperate region," *Tehnicki vjesnik-Technical Gazette*, vol. 26, no. 4, pp. 1052–1060, 2019.
- [12] A. K. S. Sabonchi and B. Akay, "Cryptanalysis of poly-alphabetic cipher using differential evolution algorithm,"



## Retraction

# Retracted: A Prediction Model Analysis of Behavior Recognition Based on Genetic Algorithm and Neural Network

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] Q. Wang and S. Liu, "A Prediction Model Analysis of Behavior Recognition Based on Genetic Algorithm and Neural Network," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3552908, 11 pages, 2022.

## Research Article

# A Prediction Model Analysis of Behavior Recognition Based on Genetic Algorithm and Neural Network

Qifu Wang<sup>1</sup> and Shuzhi Liu <sup>2</sup>

<sup>1</sup>School of Sports Science, Changsha Normal University, Changsha 410100, China

<sup>2</sup>School of Physical Education, Chongqing University, Chongqing 400044, China

Correspondence should be addressed to Shuzhi Liu; [cqutyliushuzhi@cqu.edu.cn](mailto:cqutyliushuzhi@cqu.edu.cn)

Received 14 January 2022; Revised 15 February 2022; Accepted 3 March 2022; Published 26 March 2022

Academic Editor: Daqing Gong

Copyright © 2022 Qifu Wang and Shuzhi Liu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the extensive application of virtual technology and simulation algorithm, motion behavior recognition is widely used in various fields. The original neural network algorithm cannot solve the problem of data redundancy in behavior recognition, and the global search ability is weak. Based on the above reasons, this paper proposes an algorithm based on genetic algorithm and neural network to build a prediction model of behavior recognition. Firstly, genetic algorithm is used to cluster the redundant data, so that the data are in fragment order, and then it is used to reduce the data redundancy of different behaviors and weaken the influence of dimension on behavior recognition. Then, the genetic algorithm clusters the data to form subgenetic particles with different dimensions and carries out coevolution and optimal location sharing for subgenetic particles with different dimensions. Through simulation test, the algorithm constructed in this paper is better than genetic algorithm and neural network algorithm in terms of calculation accuracy and convergence speed. Finally, the prediction model is constructed by setting the initial value and threshold to predict the behavior recognition, and the results show that the accuracy of the model constructed in this paper is improved in the analysis of behavior recognition.

## 1. Introduction

In recent years, the state has put forward the development model of behavior intelligence, and major scientific research institutions attach great importance to behavior analysis. According to the survey data of behavior intelligence in 2020 [1], behavior recognition and behavior prediction account for 45.6% of the development of intelligent technology, resulting in the increasing demand for behavior intelligence development analysis and penetration into society. China has established research institutions of artificial intelligence, artificial voice, and artificial portrait and applied behavior recognition to video monitoring to provide data for artificial intelligence analysis [2]. Some scholars believe that behavior recognition is the judgment of different behavior points. The advantage of genetic algorithm is to analyze the gradual process of behavior and judge the final result of behavior. However, genetic algorithm also has its own shortcomings, which cannot adjust the analysis depth according to the

behavior results. Neural network makes up for the deficiency of genetic algorithm. Its advantage is to carry out self-learning, analyze the relationship between different behaviors, and make the theoretical behavior consistent with the actual behavior. Compared with the general algorithm, the combination of genetic algorithm and neural network increases the coupling between different behavior nodes and makes the behavior analysis more dynamic. At the same time, the increase of coupling can reduce the computational complexity and obtain the calculation results more accurately. The combination point of genetic algorithm and neural network is the judgment of behavior nodes and coupling between nodes. Genetic algorithm gives play to its advantage of progressive analysis of behavior nodes, and neural network analyzes the coupling between nodes, so as to improve the accuracy of analysis results. Behavior recognition is a dynamic process, which involves many factors. It is generally calculated by genetic algorithm. Although the node analysis of genetic algorithm can judge different

behaviors, it lacks threshold adjustment and cannot realize the dynamic analysis of behaviors. The combination of genetic algorithm and neural network can not only judge the behavior points but also judge the relationship between different behaviors. Therefore, the research advantage of this paper is to realize the comprehensive judgment of behavior and improve the accuracy of behavior recognition. On the basis of previous studies, this paper establishes a behavior recognition prediction model to provide accurate information and technology for behavior recognition prediction in sports and enrich relevant theories.

As can be seen from Figure 1, the application of behavior recognition in the field of sports is deepening and increasing and has become the focus of the development of intelligent sports. Among them, shooting, track and field, ball games, and other fields have been improved from 2015 to 2019, but the proportion of various sports has not changed. It can be seen that the social research degree of sports behavior recognition is increasing, which indicates the importance of verification of sports behavior recognition. The influencing factors of sports behavior recognition have the characteristics of semi-structured and nonlinear relationship, and the genetic algorithm has better semi-structured data analysis ability and can analyze it well. The neural network assigns value to each influencing factor and calculates it. If the output does not meet the requirements, adjust the assignment and propose the factor with assignment  $<0.3$ . Through continuous proposal, the purpose of simplifying factors is achieved. Finally, the factor with a value  $>0.6$  is the key factor, and subsequent calculation is carried out. It uses the mapping characteristics between various factors to carry out repeated iteration and self-learning [3], better process behavior recognition data, and meet the requirements of nonlinear analysis. Some scholars have established a behavior recognition development prediction model combining Apriori algorithm and genetic algorithm [4] to analyze domestic behavior recognition. The results show that the joint analysis results of genetic algorithm and Apriori algorithm are better, and each index is significantly better than Apriori algorithm[5]. Genetic algorithm belongs to a comprehensive analysis method, which integrates the learning mechanism of neural network to cluster a series of problems to achieve the purpose of fuzzy analysis. Therefore, based on genetic algorithm, this paper predicts behavior recognition [6]. In previous studies, the neural network is improved mainly from the following aspects. (1) Parameter adjustment and learning strategy change: the development trend of behavior recognition is studied and three behavior recognition prediction methods are compared. The results show that selecting appropriate methods can solve the eigenvalues of single index and multi-index in a certain period of time [7]. (2) A collection of analysis methods with other behavior recognition: some scholars have proposed a hybrid adaptive neural network algorithm, which integrates K-means clustering into the neural network. The algorithm can classify the behavior recognition data [8], divide the neural network according to different clustering weights, and improve the search speed of the optimal value. However, this method is greatly affected by the Euclidean distance,

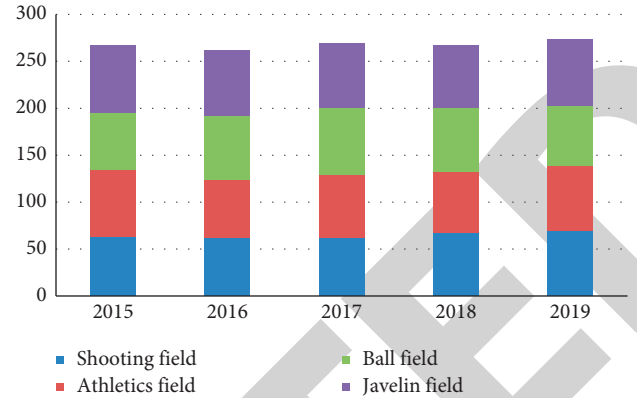


FIGURE 1: The application of behavior recognition in sports.

which is not conducive to the calculation of the global optimal value [9]. From the perspective of communication delay, the neural network algorithm is used for large-scale scheduling calculation of distributed behavior recognition, and the dynamic and random cluster analysis [10] is used to prove that the algorithm has high accuracy. (3) Compared with classical theories, such as Bayesian theory, grey theory, chaos theory, and rough set, the improved neural network method has strong time-series sensitivity to data [11]. The above research is based on this, but neural network algorithm's calculation accuracy, search time, and overall convergence are still not ideal. In this paper, K-means clustering is used to improve the neural network algorithm and then combined with genetic algorithm to build a behavior recognition prediction model and verify it.

## 2. The Description of Behavior Recognition Algorithm Based on Genetic Algorithm and Neural Network

**2.1. Behavior Identification.** Behavior recognition is to judge the behavior of natural organisms. Behavior recognition is divided into two parts. One part is behavior recognition, that is, the movement of limb behavior. The other part is psychological behavior, which is the psychology that dominates the behavior of natural organisms. Generally speaking, behavior recognition refers to limb behavior, that is, the behavior and movement of natural organisms. Behavior recognition is related to the surrounding environment, movement trend, and movement purpose. Behavior recognition is not only a combination of individual behaviors but also a judgment of future behaviors.

**2.2. The Neural Network Theory.** Neural network was first proposed by Box and Jenkins in 1927 [12]. It belongs to a statistical analysis method, which mainly solves the problems of behavior description, behavior analysis, and future prediction, and is widely used in various fields. The algorithm is superior in nature, which is reflected in the description of data persistence. The neural network randomly selects irregular points for cyclic analysis. Neural network algorithm principle and structure are as follows: let  $G_{it}$  be the

dataset of  $t$  time in  $i$  multi-dimensional space and  $x$  belong to  $G_{it}$ . Hypothesis 1:  $P_j(x) = \sum_{i,t=1}^n \{x_{it}^j \cdot k + \varphi(k)\}$ ,  $k \in (1, \dots, n)$ ; if the deviation  $\varphi(k) = C(x, \varepsilon)$  is satisfied, then  $P_j(x)$  is the best genetic set where  $C(x, \varepsilon)$  is the adjustment coefficient of  $x$  and  $\varepsilon$ , and  $\varepsilon$  is the “core point” on the right side of the neural network. When,  $x_{it} = 2 \sin(\tan \pi t/\rho)$  and  $\rho$  is the minimum prime set  $2 \sin(\tan \pi t/\rho)^j \cdot k < \sum_{i,j,t=1}^n \{x_{it}^j \cdot k$ , the data link of the neural network list is shorter.

The relevant theories of neural network are as follows.

**Theorem 1.** *If the local deviation of  $P_n(x)$  is  $\varphi(x \cdot k)$  and  $f(\cdot) \in B_t$  ( $B_t$  is the redundancy function class of  $t$  time), the calculation formula of the redundancy function is*

$$\int_{x \in G_{ij}} f(x) dx - \sum_{j=1}^n f(P_j(x)) < Q(f(x)) \cdot \varphi(k), \quad (1)$$

where  $Q(f(x))$  is the global redundancy of  $f(\cdot)$ .

**Theorem 2.** *If the following conditions are satisfied  $f(x)$ , then any function value of  $f(x) < \varepsilon$ ,  $f(x)^i < \varepsilon, \dots, f(x)^j < \varepsilon$ , on the right side of the neural network is within the integral of  $f(x)$ , and the error of the redundant data on the right side is less than  $\ln(1/x)$ .*

**Theorem 3.** *If any data point  $x_{it}$  is normally distributed on the behavior recognition axis, the calculation formula of uniform deviation of  $D(x)$  any point is as follows.*

$$D(x, f(x)) = \ln\left(\frac{1}{\sqrt{x}} \cdot \sqrt[{}^t]{\log(P_j(x))}\right), \quad (2)$$

where  $D(x, f(x))$  is between sets  $[0, 1]$ .

It can be seen from Theorems 1 and 2 that the relationship between the right redundancy amplitude and the amount of data  $x$  can be obtained by using the approximate integration of neural network, which is independent of spatial dimension  $i$  and time  $t$ . Therefore, neural network function provides a good theoretical basis for multi-dimensional number calculation and time-dependent calculation and reduces the impact of unstructured data form on the calculation results. According to Theorem 3, the local deviation of the data in the neural network is  $\ln(1/x)$  and the uniform deviation is  $\ln(1/\sqrt{x} \cdot \sqrt[{}^t]{\log(P_j(x))})$ , indicating that the local deviation can control the redundancy amplitude on the right side due to the uniformly distributed deviation. Combining neural network with rough set algorithm, an improved algorithm of right redundant amplitude control is proposed, which proves the ability of neural network to control redundant data and verifies the superiority of this method. This paper uses this method for reference to transform the composition of behavior recognition data and takes it as the initial data group for continuous analysis.

**2.3. Genetic Algorithm.** In this paper, genetic algorithm is selected for model optimization. This method belongs to a clustering method with fuzziness and adaptability. It can

realize self-learning [13] and constantly revise the membership function of cluster set and complete set. Genetic algorithm can be defined by “IF” judgment mode. Under constraint rule  $Z$ , neural network algorithm clustering process is as follows:

$$\begin{aligned} \text{IF: } x_i \in g_{ij} \text{ and } Z(x) = C(x_i, x_{i-1}), \\ \text{then } y = \sum_{i=1}^n y_i = \sum_{i=1}^n \frac{Z(x_i) \cdot \lambda(x_i)}{\lambda(x_i)}, \end{aligned} \quad (3)$$

where  $\lambda(x_i)$  is the clustering function,  $g_{ij}$  is the clustering set,  $M(x)$  is the constraint function,  $C(x_i, x_{i-1})$  is the relationship between any adjacent data  $x$ , and  $y$  is the clustering result. The input part of genetic algorithm is clear, the processing part is fuzzy, and the output part is clear, so the output result can be inferred by fuzzy method to obtain the output combination.

Assuming that any input variable is  $x_i$ , the membership relationship between each input variable  $x_i$  and output variable  $y_i$  can be obtained by constraint rule  $Z$ , as shown in the following formula:

$$C \cdot g_{ij} = \sum_{i,j=1}^n \exp\left\{\frac{(x_i - c_{ij})}{b_{ij}}\right\}. \quad (4)$$

where  $c_{ij}$  and  $b_{ij}$  are, respectively, the center and degree of membership of the membership function;  $C$  is a relational function; and  $g_{ij}$  is a fuzzy set.

By clustering the above membership relations, the continuous operator of clustering can be obtained, as shown in the following formula:

$$C \cdot g_{ij} = \delta \cdot \sum_{i,j,k=1}^n g_{ij}^k(x), \quad (5)$$

where  $\delta$  is the cluster adjustment coefficient and  $k$  is the derivation coefficient. According to the cluster calculation, the output value of the cluster can be obtained, as shown in the following formula:

$$y = \frac{[\delta \cdot \sum_{i,j,k=1}^n g_{ij}^k(x) \cdot f(P_j(x))]}{\max \sum_{i,j=1}^n g_{ij}^k(x)}. \quad (6)$$

Genetic algorithm not only shortens the processing time of behavior recognition data but also increases the amount of initial data processing. Genetic algorithm performs multi-dimensional clustering according to the initial data volume [14] and labels it with time to form a continuous data chain. Adjust the relationship function  $C$  and clustering coefficient of  $\delta$  the data according to the output result  $y$ .

### 3. The Construction of Behavior Recognition Model Based on Genetic Algorithm and Neural Network Algorithm

**3.1. Coevolution of Initial Behavior Identification Data.** In this paper, the model takes into account the multi-dimensional and global nature of the data, reduces the local convergence, improves the search ability of the characteristic

data, and shortens neural network algorithm's own learning time. Reflected in the evolution strategy of initial sequence and multi-dimensional sequence, the convergence threshold and weighting factor are used to promote the coevolution of distributed population, and the optimal value of right redundancy amplitude is obtained.

**3.1.1. Initialization of Genetic Algorithm.** The neural network algorithm considers that the initial sequence is in random distribution. If the sequence uncertainty is very strong, that is, nonnormal distribution, it will make the sequence calculation fall into local optimization, reduce the calculation time of global optimization solution, and increase the error rate of calculation results. In order to enhance the normality of the sequence, we should expand the scale of the sequence and increase the types of data. Based on the principle of classical neural network, this paper expands the number of sequences of initial data and improves the diversity of sequences.

Figures 2 and 3 show multi-dimensional initial neural networks processed by a random method and the genetic algorithm method, respectively. The number of segments of the neural network is 100. Through comparison, it is found that the initial sequence generated by random method is more chaotic and nondirectional. Through the comparison of Figures 2 and 3, it can be seen that the behavior recognition particles in random mode are more chaotic and concentrated, which increases the complexity of calculation. However, the behavior recognition particles processed by genetic algorithm are more uniform, which reduces the computational complexity and lays the foundation for later calculation. The sequences processed by genetic algorithm are more concentrated and directional. According to Theorems 1 and 2 of neural network algorithm, the dataset constructed by this algorithm is independent of spatial dimension and is suitable for multi-dimensional data processing. Moreover, each time the point is taken, the distribution effect of the initial sequence is the same, and the stability of the dataset is high. To sum up, genetic algorithm is selected to process the initial sequence.

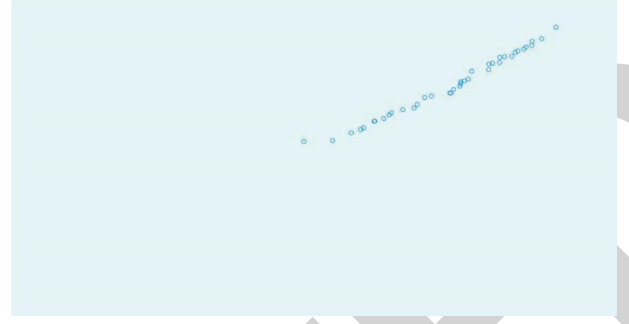


FIGURE 2: Initial neural network constructed by random method.

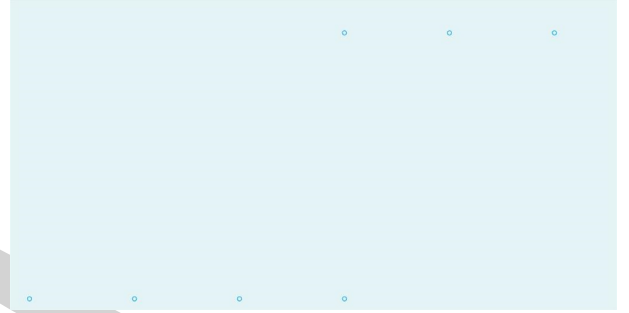


FIGURE 3: Initial neural network constructed by genetic algorithm.

**3.1.2. Multi-Strategy Collaboration.** Global search ability is an important measure of the algorithm. In the initial stage of calculation, the algorithm focuses on global search and then advances to local search [15]. In order to improve the computational efficiency of the algorithm, this paper introduces a multi-strategy cooperation method, which allows particles to search according to different behavior sequences (saving repeated calculation time), and realizes global search and local search at the same time. At present, there are standard neural network algorithms and other improved algorithms.

(1) Classical behavior recognition model:

$$y_{ij}(t+1) = \square \bar{w} \cdot y_{ij}(t) + c_1 \cdot r_1 \cdot \frac{[\delta \cdot \sum_{i,j,k=1}^n g_{ij}^k \{x(t) \cdot f(P_j[x(t)])\}]}{\max \sum_{i,j=1}^n g_{ij}^k [x(t)]} \quad (7)$$

(2) Group behavior recognition model:

$$y_{ij}(t+1) = \omega \cdot y_{ij}(t) + c_2 \cdot r_2 \cdot \frac{[\delta \cdot \sum_{i,j,k=1}^n g_{ij}^k \{x(t) \cdot f(P_j[x(t)])\}]}{\max \sum_{i,j=1}^n g_{ij}^k [x(t)]} \quad (8)$$

(3) Behavior recognition judgment model:



$$y_{ij}(t+1) = n \cdot \left| \frac{\left[ \delta \cdot \sum_{i,j,k=1}^n g_{ij}^k \{x(t) \cdot f(P_j[x(t)])\} \right]}{\max \sum_{i,j=1}^n g_{ij}^k [x(t)]} - \frac{\left[ \vec{\delta} \cdot \sum_{i,j,k=1}^n g_{ij}^k \{x(t) \cdot f(P_j[x(t)])\} \right]}{\max \sum_{i,j=1}^n g_{ij}^k [x(t)]} \right| \quad (9)$$

(4) General behavior recognition model:

$$y_{ij}(t+1) = \omega \cdot y_{ij}(t) + \sum_{i=1}^2 c_i \cdot r_i \cdot \frac{\left[ \delta \cdot \sum_{i,j,k=1}^n g_{ij}^k \{x(t) \cdot f(P_j[x(t)])\} \right]}{\text{avg} \sum_{i,j=1}^n g_{ij}^k [x(t)]}, \quad (10)$$

where  $c_1$  and  $c_2$  are the synergy coefficients of different strategies,  $r_1$  and  $r_2$  are the correlation coefficients of different strategies,  $\omega$  is the average weight of different strategies ( the values are from relevant literature at home and abroad ), and  $n$  is the number of different strategies.

This paper improves the standard neural network function in two aspects. On the one hand, it expands the search range. For each iterative calculation, a neural network form will be randomly selected from the above 5 forms for genetic algorithm calculation. Genetic algorithm can avoid local optimal results and expand the search range. The random selection calculation form can maintain the population diversity and increase the possibility of obtaining the local optimal value. On the other hand, it can improve the convergence. In order to balance the ability of global search and local search of neural network,  $\alpha$  (nonlinear adjustment coefficient) and  $\omega$  (linear weight) are added to improve the calculation speed. The formula is as follows:

$$\alpha = \frac{\text{Line}_t^d - 1}{\text{Line}_T^D}, \quad (11)$$

where  $e$  is the logarithmic base,  $t$  is the time,  $T$  is the maximum time,  $d$  is the number of iterations, and  $D$  is the maximum number of iterations. In the initial stage, the attenuation degree of  $\alpha$  is low, and the global optimal value can be calculated. In the middle and later stages of  $\alpha$  calculation, the attenuation degree increases gradually, and the local optimal value can be calculated. The linear weight is calculated as follows:

$$w = \sum_{d=1}^D \left[ \frac{(\widetilde{w_{\max}} - \sum_{t=1}^T \Delta w_i)}{(\sum_{t=1}^T \Delta w_i - w_{\min})} \right]^d, \quad (12)$$

where  $w_{\max}$  and  $w_{\min}$  are the maximum and minimum weights, respectively, and the values of  $d$ ,  $t$ ,  $D$ , and  $T$  are the same as above.

**3.1.3. The Behavior Recognition and Collaboration Strategies among Different Dimensions.** The algorithm adopts differentiated operations and strategies for data of different dimensions and adjusts corresponding parameters to realize multi-dimensional distributed collaboration, so as to

complete the process of collaborative evolution. In the model, the neural network data are divided into five dimensions, and each dimension represents the subspace of the solution space [16]. In each iteration, the five dimensions evolve simultaneously. After an iterative calculation, compare the fitness values of different dimensions and record the location of the global optimal value of each dimension. Then, each subdimension learns gradually to the global optimal value and obtains the optimal position of the subdimension in the most concise way, so as to improve the speed and efficiency of search and calculation.

**3.2. The Behavior Recognition and Judgment Method Based on Genetic Algorithm and Neural Network.** The basic idea of genetic algorithm is multi-dimensional coevolution. The initial value and threshold of neural network (the values are from relevant literature at home and abroad) are adjusted and optimized [17] to obtain the optimal solution and reduce the redundancy rate in behavior recognition.

The judgment process is shown in Figure 4.

**Step 1.** Determine the dimension and sequence structure of the neural network and determine the data structure of the neural network according to the data characteristics and problem-solving needs. The initial weight and threshold of the whole data are mapped to the neural network as a whole. The dimension of each neural network is the opportunity of weight and threshold. According to the actual application, the dimension of neural network in this paper is  $g = 433$ .

**Step 2.** Data initialization: the relevant parameters of the neural network are initialized randomly. Let the number of neural networks be  $n = 100$ , the maximum weight be  $w_{\max} = 0.7$ , the minimum weight be  $w_{\min} = 0.3$ , the synergy constant between different dimensions be  $c_1, c_2 = 3.22$ , and the maximum number of iterations be  $D = 10$ .

**Step 3.** Generate the appropriate function. The genetic algorithm theory is used to generate the initial sequence of neural network, which is mapped to the neural network as the initial weight and threshold. Through the self-learning and training of formulas (1)–(8), the synergy coefficient of the neural network algorithm will continue to increase and adapt. The accuracy of each neural network is calculated, and



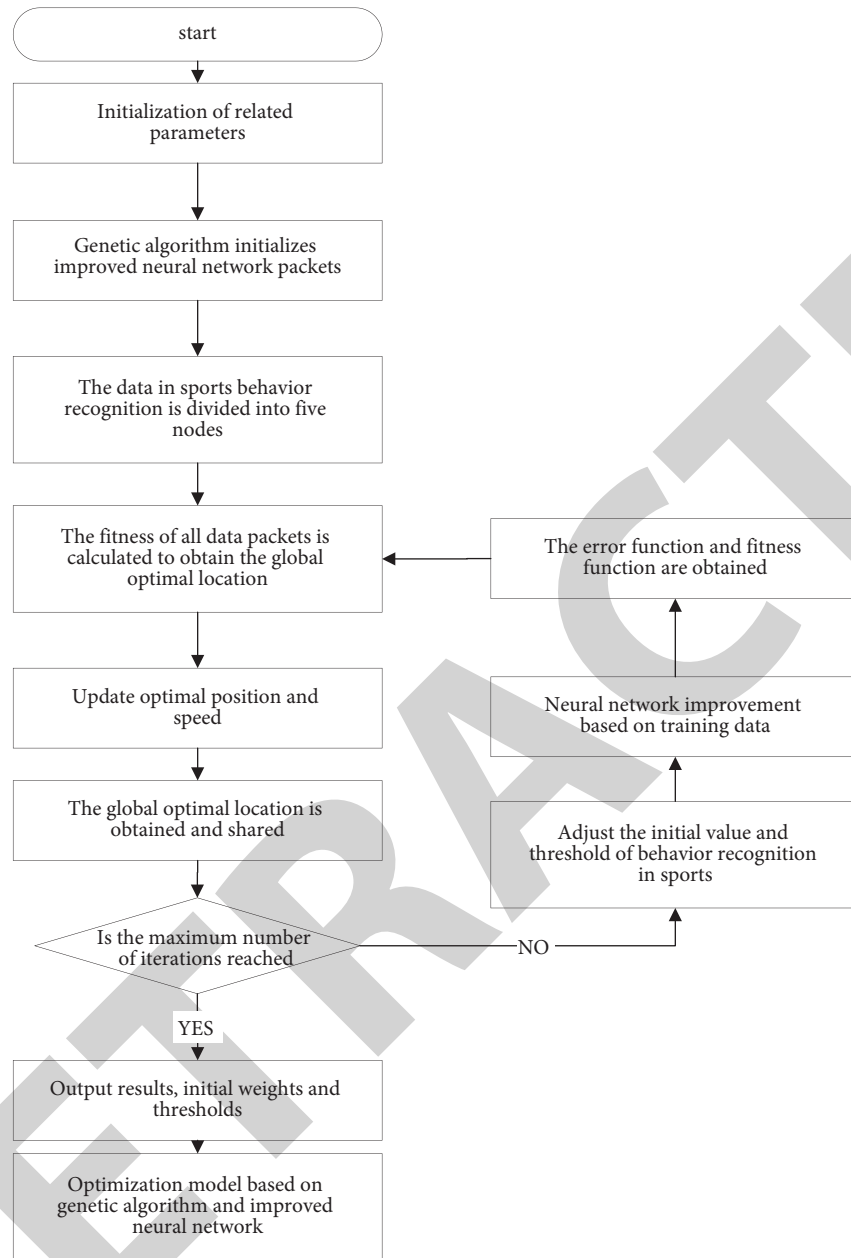


FIGURE 4: Calculation flowchart based on genetic algorithm and neural network.

the absolute value of the sum of squares is taken as an appropriate function.

*Step 4.* Globally search the optimal position of the neural network and the optimal position of each subgenetic particle. The initial neural network is randomly divided into five subsequences, the fitness ratio is obtained, and the optimal global position and the optimal position of each subsequence are recorded.

*Step 5.* Iteration of optimal position and speed: in the evolution of 5 seed sequences, one is randomly selected for evolution, such as formulas (8)–(11), and  $\alpha$  is nonlinear

adjustment coefficient and  $\omega$  is linear weight, such as formulas (12) and (13).

*Step 6.* All subsequences coevolute at the same time. After one iteration, the optimal global computing location is selected and shared with other subsequences. Other subsequences gradually iterate to this position to obtain the best position result.

*Step 7.* Determine whether to stop the iteration. If the number of iterations  $D$  is less than the maximum number of iterations  $D$ , repeat Steps 2–6; otherwise, stop the iteration and return the results such as threshold, weight, and optimal position.

TABLE 1: The test results of different functions.

Detection function	Algorithm	Value range	Ad (E)	Sd	Global optimal solution
Filer	Genetic algorithm	1.81E-6~0.13e-11	1.33-7	0.91 E-6	0
	Neural network algorithm	1.61E-7~0.04e-10	1.13-8	0.31 E-8	
	Genetic algorithm and neural network algorithm	1.43E-9~0.13e-11	0.91-9	0.17 E-10	
Rastring	Genetic algorithm	3.33E-7~1.13e-11	1.01-7	0.90 E-8	0
	Neural network algorithm	1.63E-4~1.43e-10	1.11-4	1.13 E-5	
	Genetic algorithm and neural network algorithm	1.31E-7~1.81e-11	1.14-7	1.37 E-8	
Ackley	Genetic algorithm	6.71E-7~1.13e-13	4.11-7	1.11 E-8	0
	Neural network algorithm	4.71E-6~1.13e-13	1.13-6	0.77 E-7	
	Genetic algorithm and neural network algorithm	4.53E-6~1.13e-11	1.35-6	0.11 E-7	

#### 4. Case Analysis of Sports Behavior Identification

*4.1. Model Performance Judgment.* The classical neural network algorithm, neural network algorithm, and neural network algorithm based on genetic algorithm which are tested with Filer, Rastring, Ackley, and other functions to verify the performance of the model proposed in this paper.

Filer function is the globally unique minimum value function of the test model [18], and the formula is as follows:

$$F_1(x) = \sum_{i=1}^n [x_i^2]. \quad (13)$$

Rastring function generates local minimum frequently through cosine modulation transfer function to verify the practicability of model solution [19]. The formula is as follows:

$$F_2(x) = \sum_{i=1}^n [x_i^2 - 10 \cos(2\pi x_i) + 10]. \quad (14)$$

Ackley function is a gradient optimization function of multi-dimensional points to test the calculation speed of multi-dimensional data to detect the global convergence speed. The formula is as follows:

$$F_3(x) = -20 \exp \left\{ -0.2 \sum_{i=1}^n \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2} \right\} - \exp \left\{ \frac{1}{n} \sum_{i=1}^n \cos(2\pi x_i) \right\} + 20 + e, \quad (15)$$

where  $n$  is the total number of indicators for calculating data and  $x_i$  is any number of indicators. The value range of  $x_i$  in filer function is  $\{-100100\}$ , that in rastring function is  $\{-5.12, 5.12\}$ , and that in Ackley function is  $\{-32, 32\}$ .

In order to facilitate calculation, the number of neural networks in this paper is  $n = 100$ , the maximum number of iterations is  $D = 30$ , and the maximum time is  $T = 16$  months. The above three functions are tested, respectively. In order to reduce the impact of randomness on the results, take the average of 7 times the results. The specific calculation results are shown in Table 1.

The convergence diagram of each data in Table 1 is shown in Figures 5-7.

It can be seen from Table 1 that compared with neural network and genetic algorithm, the genetic algorithm and neural network algorithm proposed in this paper is closer to the global optimal value. In terms of standard deviation, average value, and value range, genetic algorithm and neural network algorithm are better than the other two algorithms. According to the curve changes in Figures 5-7, the genetic and neural network algorithms have better stability and faster convergence speed. Therefore, the convergence speed, calculation accuracy, and summation stability of genetic and neural network algorithms are better.

*4.2. Case Data and Preprocessing.* There are some missing data in the sample data in this paper, which is compensated by default value in the later stage. The prediction dataset of behavior recognition includes digital behavior, bionic behavior, natural behavior, psychological behavior, and expected behavior. After preliminary data preprocessing, 1982 lines of structured data and 38 lines of semistructured data are obtained. In order to facilitate data analysis, behavior recognition is divided into five levels: relatively intelligent, intelligent, general, not intelligent, and relatively not intelligent. The processing results of data volume are shown in Table 2.

Due to lack of some data in this paper, the data from January 1, 2019, to July 22, 2020, are taken as the training set, with a total of 1345 pieces; the data from July 22, 2020, to December 1, 2021, are taken as the test set, with a total of 676 pieces.

*4.3. Test Results.* In order to verify the genetic and neural network algorithms proposed in this paper, the results are compared with those of genetic algorithm and neural network algorithm. The results are shown in Figure 8.

As can be seen from Figure 8, in terms of calculation accuracy, The accuracy of the algorithm proposed in this

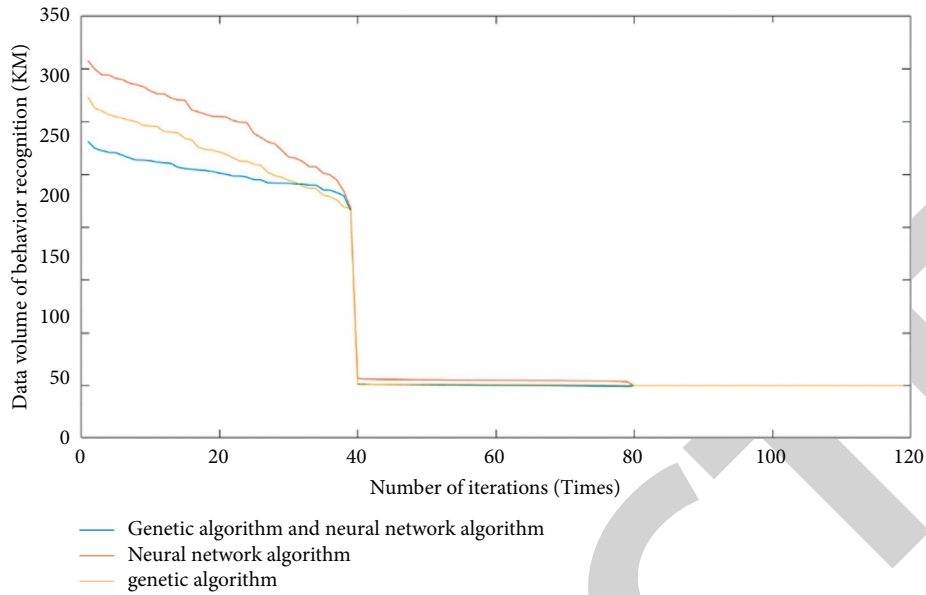


FIGURE 5: The optimization convergence result of filer function.

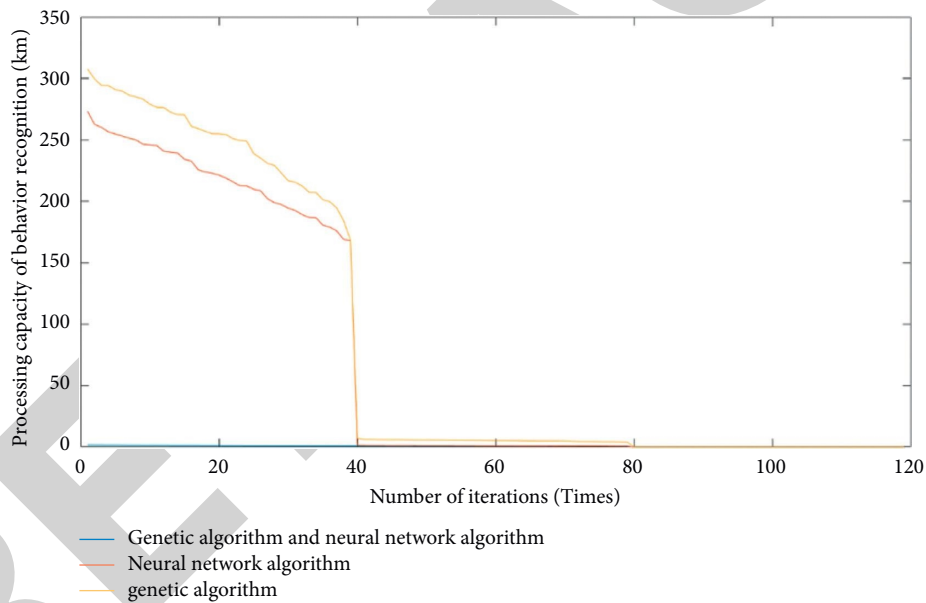


FIGURE 6: The convergence results of rastring function optimization.

paper is higher than the single genetic algorithm and neural network algorithm, respectively, after analysis the error rate is also lower than that of the single genetic algorithm and neural network algorithm. In addition, the algorithm proposed in this paper has better computational stability, which is better than the single genetic algorithm and neural network algorithm. The calculation results of different algorithms are shown in Table 3.

It can be seen from Table 3 that the single neural network algorithm and genetic algorithm have the problems of insufficient accuracy and large variation of calculation results in different levels of behavior recognition and prediction. Compared with the neural network algorithm, the accuracy

of the algorithm constructed in this paper is significantly improved. At the same time, the accuracy is close to that of neural network algorithm, which is more than 80%, which is better than genetic algorithm. In order to further verify the advantages of genetic and neural network algorithms, the optimal fitness values of different algorithms are compared, and the results are shown in Figure 9.

It can be seen from Figure 9 that the optimization of fitness function results of genetic and neural network algorithms is more significant. The reason is that genetic algorithm and neural network add synergy coefficients, improvement weights, and convergence factors in different dimensions.

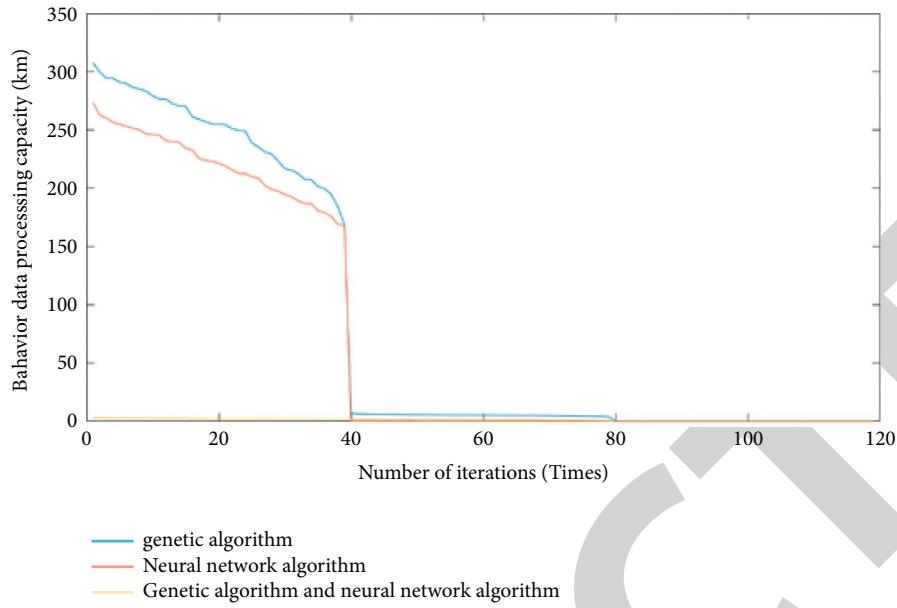


FIGURE 7: The convergence results of Ackley function optimization.

TABLE 2: The classification and proportion of sports behavior recognition.

Identification level	Data volume (PCs.)	Proportion (%)
Intelligent identification	483	23.90
Accurate identification	1043	51.61
General identification	332	16.43
Unidentification	97	4.80
Special identification	66	3.27

Note. The indicators of this paper come from domestic literature and expert questionnaire.

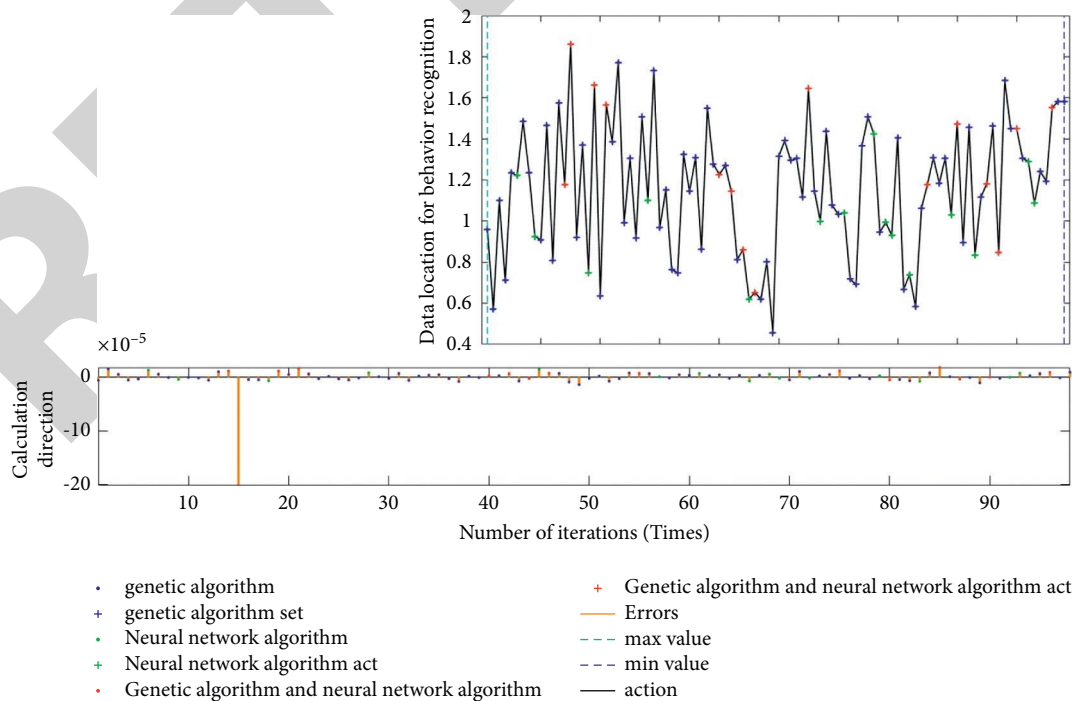


FIGURE 8: Test results of different algorithms.

TABLE 3: The comparison of prediction accuracy of different grades.

Algorithm	Intelligent identification	Accurate identification	General identification	Unidentification	None identification
Genetic algorithm	47.23	59.22	37.28	38.34	35.94
Neural network algorithm	82.32	80.13	81.33	82.47	81.91
Genetic algorithm and neural network	86.41	86.27	87.17	85.49	87.70

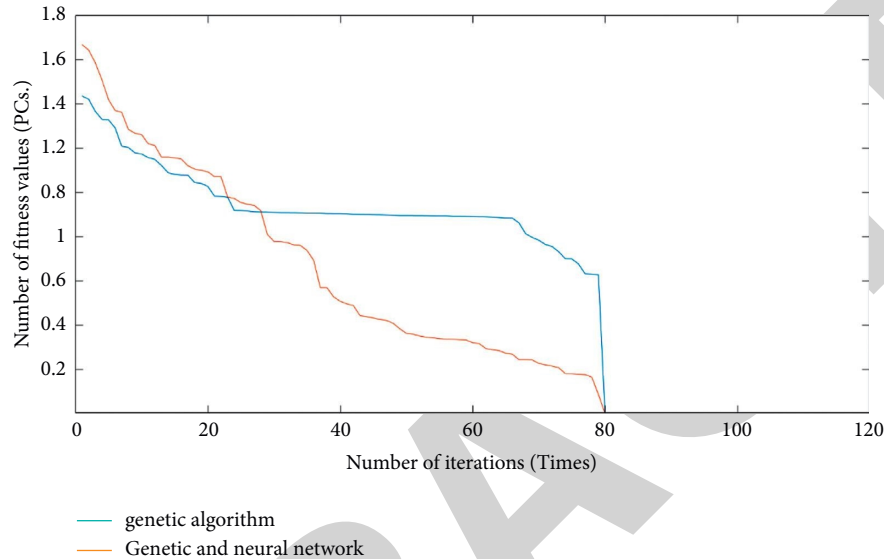


FIGURE 9: Fitness optimization results of genetic algorithm and combined genetic and neural network algorithms.

## 5. Conclusion

To sum up, the model of genetic algorithm combined with neural network constructed in this paper can identify and predict behavior more accurately and quickly, and its calculation results are basically consistent with the actual requirements. This paper puts forward the theory of genetic algorithm and improves the neural network algorithm combined with multi-dimensional coevolution method. At the same time, the threshold and weight of the neural network algorithm are set, and the behavior recognition analysis model is constructed. The results show that compared with the classical neural network algorithm and genetic algorithm, the prediction accuracy and convergence of behavior recognition based on genetic algorithm combined with neural network algorithm are better, and the behavior recognition can be judged. However, in this model, the multi-dimensional cooperation strategy pays too much attention to the global search ability, resulting in the relative decline of the local search ability and the reduction of the calculation speed of the optimal solution. Therefore, in the future research, the equilibrium coefficient and filter function will be added to improve the solving ability of the model.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This study was supported by the <https://doi.org/10.13039/501100012226> Fundamental Research Funds for the Central Universities in <https://doi.org/10.13039/501100002369> Chongqing University (2021CDSKXYTY003). The authors are grateful for the support.

## References

- [1] H. Cao, H. Nguyen, T. Tran, H. Tran, and J. Jeon, "A robot calibration method using a neural network based on a butterfly and flower pollination algorithm," *IEEE Transactions on Industrial Electronics*, vol. 69, no. 4, pp. 3865–3875, 2022.
- [2] Y. J. Xu, F. Li, and A. Asgari, "Prediction and optimization of heating and cooling loads in a residential building based on multi-layer perceptron neural network and different optimization algorithms," *Energy*, vol. 14, no. 32, p. 240, 2022.
- [3] Q. Ye, L. Zhenghao, Z. Haoxin, and Z. Yongmei, "Human behavior recognition based on time correlation sampling two-stream heterogeneous grafting network," *Optik*, vol. 5, no. 13, p. 251, 2022.
- [4] S. D. Chen, X. Qingwen, Z. Xiaochen, X. Yingying, and L. Jian John, "Risky driving behavior recognition based on vehicle trajectory," *International Journal of Environmental Research and Public Health*, vol. 18, no. 23, p. 45, 2022.

## Retraction

# Retracted: Differentially Private Singular Value Decomposition for Training Support Vector Machines

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] Z. Sun, J. Yang, and X. Li, "Differentially Private Singular Value Decomposition for Training Support Vector Machines," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 2935975, 11 pages, 2022.



## Research Article

# Differentially Private Singular Value Decomposition for Training Support Vector Machines

Zhenlong Sun <sup>1,2</sup>, Jing Yang <sup>1</sup> and Xiaoye Li <sup>2</sup>

<sup>1</sup>College of Computer Science and Technology, Harbin Engineering University, Harbin 150001, China

<sup>2</sup>College of Computer and Control Engineering, Qiqihar University, Qiqihar 161006, China

Correspondence should be addressed to Jing Yang; yangjing@hrbeu.edu.cn

Received 25 January 2022; Revised 16 February 2022; Accepted 21 February 2022; Published 26 March 2022

Academic Editor: Daqing Gong

Copyright © 2022 Zhenlong Sun et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Support vector machine (SVM) is an efficient classification method in machine learning. The traditional classification model of SVMs may pose a great threat to personal privacy, when sensitive information is included in the training datasets. Principal component analysis (PCA) can project instances into a low-dimensional subspace while capturing the variance of the matrix  $A$  as much as possible. There are two common algorithms that PCA uses to perform the principal component analysis, eigenvalue decomposition (EVD) and singular value decomposition (SVD). The main advantage of SVD compared with EVD is that it does not need to compute the matrix of covariance. This study presents a new differentially private SVD algorithm (DPSVD) to prevent the privacy leak of SVM classifiers. The DPSVD generates a set of private singular vectors that the projected instances in the singular subspace can be directly used to train SVM while not disclosing privacy of the original instances. After proving that the DPSVD satisfies differential privacy in theory, several experiments were carried out. The experimental results confirm that our method achieved higher accuracy and better stability on different real datasets, compared with other existing private PCA algorithms used to train SVM.

## 1. Introduction

In the past decade, more and more personal information has been stored in electronic databases for machine learning and personalized recommendation. The data sharing and analysis bring lots of convenience to people's lives, but pose a great threat to personal privacy. Support vector machine (SVM) [1] is a popular classification method that searches for the best hyperplane that separates two class instances by solving a quadratic optimization problem. It has been applied in pattern recognition such as image recognition and text classification. In the classification model of SVM, the most serious privacy issue is that the support vectors (SVs) are directly obtained from the training datasets [2]. Therefore, the classification model should be privately published to avoid disclosing personal sensitive information.

Differential privacy (DP) [3–6] has a strict mathematical definition and the level of privacy protection can be quantified by a small parameter  $\epsilon$  named privacy budget. DP

has been becoming an accept standard. It guarantees that the result of an analysis is virtually independent of the addition or removal of one record. DP has attracted a growing research attention [7]. The common mechanisms for implementing DP include Laplace mechanism [8], Gaussian mechanism [9], and exponential mechanism [10].

Principal component analysis (PCA) [9] solves a low-rank subspace, which completely captures the variance of matrix  $A$ . The main advantages of working with the low-rank approximation of  $A$  include higher time and space efficiency, less noise, and removal of correlation between features. Through PCA, the original instances are projected into a low-dimensional subspace and the features become linearly independent. Eigenvalue decomposition (EVD) and singular value decomposition (SVD) are two common algorithms to perform PCA. They are related to the familiar theory of matrix diagonalization. The EVD is used for a symmetric matrix and SVD for an arbitrary matrix. Furthermore, SVD does not need to compute the matrix of covariance compared with EVD [11].

This study researches the privacy leakage problem of SVM classifier. To overcome some shortcomings in the existing private SVMs, a differentially private singular value decomposition (DPSVD) algorithm is proposed to keep SVs private in the classification model of SVM. This study makes the following innovations:

- (i) The authors proposed an idea that projecting the training instances into the low-dimensional singular subspace and the SVM can train the classification model on it while not violating the privacy requirements for the training data.
- (ii) The projection process of the DPSVD satisfies DP, and the generated singular vectors are also private which can be provided directly to users for classification testing.
- (iii) In the DPSVD, the projection process is implemented by SVD. The main advantage is that SVD does not need to calculate the matrix of covariance compared with EVD. It takes up a lot of memory space for high-dimensional data.
- (iv) Our method protects privacy of the training instances before training the classification model; many optimization methods of SVMs can be applied directly to the training progress.
- (v) After proving that the DPSVD satisfies differential privacy in theory, several experiments were carried out. The experimental results confirm that our method achieved higher accuracy and better stability on different real datasets, compared with other existing private PCA algorithms used to train SVM.

## 2. Related Work

From a privacy perspective, SVMs have serious privacy issues, because SVs tend to be directly obtained from the training datasets. There is a lot of work to solve this privacy problem based on DP. Chaudhuri et al. [12, 13] proposed two perturbation-based methods for problems like linear SVM classification. For nonlinear kernel SVM, they derived the kernel function through random projection and linearized the function. However, it is hard to analyze the sensitivity of the output perturbation, and the differentiability criteria are required in the loss function of objective perturbation. To learn SVM privately, Rubinstein et al. [14] developed two feature mapping methods by adding noise to the output classifier. But their methods only apply to the kernels that do not change with translation. Li et al. [15] designed a mixed SVM, which alleviate much of the noise through Fourier transform based on a few open-consented information. Zhang et al. [16] proposed DPSVMDVP by adding Laplace noise to the dual variables based on the error rate. Liu et al. [17] presented an innovative private classifier called LabSam by random sampling under the exponential mechanism. Sun et al. [18] proposed the DPWSS, which introduces randomness into SVM training; they also proposed another private SVM algorithm DPKSVMEL based on

exponential and Laplace hybrid mechanism [19] for the kernel SVM to prevent privacy leakage of the SVs.

PCA constructs a set of new features to describe the instances in a low-dimensional subspace. When the generated projection vectors are private, the new instances in the low-dimensional subspace are private as well, and they can be used directly to train SVMs without compromising the privacy of the instances. There are several researches on private PCA. Blum et al. [20] developed SuLQ by disturbing the matrix of covariance with Gaussian noise. However, the greatest eigenvalue might be not real, due to the asymmetry of the noise matrix. Chaudhuri et al. [21] modified SuLQ framework with a symmetric noise matrix and used it for data publishing. Dwork et al. [9] disturbed the matrix of covariance with Gaussian noise. Imtiaz and Sarwate [22, 23] and Jiang et al. [24] disturbed the matrix of covariance with Wishart noise and it guarantees that the perturbed matrix of covariance is positive semidefinite. Xu et al. [25] and Huang et al. [26] added symmetric Laplace noise to the matrix of covariance. Those methods above all generate the perturbed matrix of covariance by adding a noise matrix and then perform EVD to implement PCA. And only [26] measured the availability of the private PCA by SVM, but not to research private PCA from the privacy perspective of SVM. Recently, SVD has been widely used in collaborative filtering [27], deep learning [28], data compression [29, 30], and images watermarking [31]. There are few researches on privacy-preserving data mining based on SVD. Keyvanpour et al. [32] defined a method that combined SVD and feature selection to benefit from the advantages of both domains. Li et al. [33] gave a new algorithm for protecting privacy based on nonnegative matrix factorization and SVD. Kousika et al. [34] proposed a methodology based on SVD and 3D rotation data perturbation for preserving privacy of data.

## 3. Background

Table 1 summarizes the symbols used in this study.

*3.1. Support Vector Machines.* Given training instances  $x_i \in R^d$  and labels  $y_i \in \{1, -1\}$  the classification model of SVM can be obtained by solving the following optimization problem [35]:

$$\min_{\alpha} f(\alpha) = \frac{1}{2} \alpha^T Q \alpha - e^T \alpha. \quad (1)$$

$$\text{Subject to } \begin{aligned} 0 \leq \alpha_i \leq C, i = 1, \dots, l \\ y^T \alpha = 0, \end{aligned} \quad (2)$$

where  $\alpha$  is a dual vector;  $Q$  is a symmetric matrix,  $Q_{ij} = y_i y_j K(x_i, x_j)$ , and  $K$  is the kernel function.

Let  $x$  be a new instance. The label of  $x$  can be predicted by the decision function as follows:

$$f(x) = \sum_{i=1}^n \alpha_i y_i K(x_i, x) + b. \quad (3)$$

TABLE 1: Symbols.

Symbol	Description
$D, D'$	The adjacent matrix of datasets
$A, \bar{A}$	Matrix of covariance
$x_i \in \mathbb{R}^d$ .	Train instance
$y_i \in \{1, -1\}$ .	Label
$A$	Dual vector
$Q$	Symmetric matrix for kernel function
$K$	Kernel function
$E$	Vector composed entirely of ones
$C$	Upper limit of $\alpha$
$\lambda_i$	Eigenvalue
$v_i$ .	Eigenvector
$\Gamma$	The accumulative contribution rate of principal components
$U, V$	The singular vectors or eigenvectors matrix
$\Sigma, S$	The singular values or eigenvalues diagonal matrix
$\sigma_i$	Singular value
$I$	Unit diagonal matrix
$M$	A randomized mechanism
$O$	All subsets of possible outcomes of mechanism $M$
$\epsilon$	Privacy budget
$\beta, \delta$	Privacy parameter
$S_1, S_2$	The $\ell_1$ and $\ell_2$ sensitivity of function
$Laplace(b)$	Laplace noise (mean: 0; scale: $b$ )
$N(0, \tau^2)$	Gaussian noise with (mean: 0; deviation: $\tau$ )

In the classification model, only the SVs determine the maximal margin and correspond to the nonzero  $\alpha_i$ s, and others equal zero. From a privacy perspective, the classification model has serious privacy issues as the SVs are intact instances.

**3.2. Principal Component Analysis.** PCA computes a low-rank subspace and achieves the dimensionality reduction for high-dimensional data, shedding light on the use of private SVM in high-dimensional data classification. For a given data matrix  $D \in \mathbb{R}^{n \times d}$  with  $d$  features of  $n$  instances, the  $i$ -th row of  $D$  is denoted by  $x_i$ , and assumes that its  $\ell_2$  norm satisfies  $\|x_i\|_2 \leq 1$ . After the matrix is centralized by column, the matrix of covariance can be obtained as

$$A = \frac{1}{n} D^T D = \frac{1}{n} \sum_{i=1}^n x_i^T x_i. \quad (4)$$

The matrix of covariance is a real symmetric matrix; therefore its eigenvalues and corresponding eigenvectors can be obtained by EVD:

$$A v_i = \lambda_i v_i, \quad (5)$$

where  $\lambda_i$  is one of the eigenvalues and  $v_i$  is its corresponding eigenvector. The  $\lambda_i$  could be treated as variance of the  $i$ -th

principal component to denote its importance and is sorted in descending order. Generally, the threshold of the accumulative contribution rate of principal components  $\gamma$  ( $0 \leq \gamma \leq 1$ ) is set to decide the target dimension  $k$  by

$$\frac{\sum_{i=1}^k \lambda_i}{\sum_{i=1}^d \lambda_i} \geq \gamma. \quad (6)$$

According to the diagonalization theory of matrix and (5), it obtains another representation of EVD as follows:

$$A = V \Sigma V^T, \quad (7)$$

where  $V$  is an orthogonal matrix consisting of eigenvectors in columns and  $\Sigma$  is a diagonal matrix taking eigenvalues as diagonal entries. Compared with EVD, SVD can be applied to an arbitrary real matrix and it does not need to calculate matrix of covariance. The representation of SVD is shown as follows:

$$D = U S V^T, \quad (8)$$

where  $U$  and  $V$  are the left and right singular matrices, which consist of left and right singular vectors, respectively;  $S$  is a diagonal matrix taking singular values as diagonal entries. The singular value  $\sigma_i$  is also sorted in descending order. The relationship between EVD and SVD is as follows:

$$A = \frac{1}{n}D^T D = \frac{1}{n}(USV^T)^T(USV^T) = \frac{1}{n}VSU^T USV^T = \frac{1}{n}VSSV^T = \frac{1}{n}VS^2V^T, \quad (9)$$

$$\tilde{A} = \frac{1}{n}DD^T = \frac{1}{n}(USV^T)(USV^T)^T = \frac{1}{n}USV^T VSU^T = \frac{1}{n}USSU^T = \frac{1}{n}US^2U^T, \quad (10)$$

where  $U^T U = I$  and  $V^T V = I$ , because  $U$  and  $V$  are both made up of unit orthogonal vectors; they also are called orthonormal basis matrices. The coefficient  $1/n$  has nothing to do with the eigenvectors and the proportionality of eigenvalues. We generally use  $D^T D$  to approximate the matrix of covariance. From (9) and (10), we can conclude that the SVD of an arbitrary real matrix yields a similar result to the EVD of its matrix of covariance. In the SVD of  $D$ , the right singular vectors serve as the eigenvectors of  $D^T D$ , and the left ones serve as those of  $DD^T$ . The singular values equal the square roots of the nonzero eigenvalues of  $D^T D$  and  $DD^T$ .

### 3.3. Differential Privacy

*Definition 1* (differential privacy (see [3])). A stochastic mechanism  $M$  satisfies  $(\epsilon, \delta)$ -differential privacy, provided that, for every two adjacent matrices  $D$  and  $D'$  differing in exactly one row, and for all subsets of probable outcomes  $O \subseteq \text{Range}(M)$ ,

$$\Pr[M(D) \in O] \leq \exp(\epsilon) \times \Pr[M(D') \in O] + \delta. \quad (11)$$

When  $\delta$  equals zero,  $M$  satisfies  $\epsilon$ -differential privacy.

*Definition 2* (sensitivity (see [3])). For a given function  $q: D \rightarrow R^d$ , and adjacent matrices  $D$  and  $D'$ , the sensitivities  $S_1$  and  $S_2$  of function  $q$  can be, respectively, expressed as

$$S_1 = \max_{D, D'} \|q(D) - q(D')\|_1 \text{ and } S_2 = \max_{D, D'} \|q(D) - q(D')\|_2. \quad (12)$$

$S_1$  corresponding to  $\ell_1$  norm is usually used in Laplace mechanism, while  $S_2$  corresponding to  $\ell_2$  norm is used in Gaussian mechanism.

*Definition 3* (Laplace mechanism (see [8])). For a numeric function  $q: D \rightarrow R^d$ , with scale factor  $b = S_1/\epsilon$ . The Laplace mechanism, which adds independent random noise distributed as  $\text{Laplace}(b)$  to each output of  $q(D)$ , ensures  $\epsilon$ -differential privacy.

*Definition 4* (Gaussian mechanism (see [9])). For a numeric function  $q: D \rightarrow R^d$ , let  $\beta = S_2 \sqrt{2 \ln(1.25/\delta)}/\epsilon$ . The Gaussian mechanism, which adds independent random noise distributed as  $N(0, \beta^2)$  to each output of  $q(D)$ , ensures  $(\epsilon, \delta)$ -differential privacy.

## 4. Materials and Methods

To overcome the shortcomings in the existing private SVMs, we proposed the DPSVD. The DPSVD privately projects the

original instances to a low-dimension singular subspace and trains a SVM classification model in it to protect the privacy of training instances.

*4.1. Algorithm Description: Algorithm 1 Is the Pseudocode of the DPSVD.* The algorithm 1 describes the implementation process of the DPSVD for training a private classification model of SVM. Firstly, it generates a noise matrix sampled from Gaussian distribution, and this step does not need to symmetrize the noise matrix as the existing private PCA algorithms. Secondly, it adds the noise matrix to the raw data matrix rather than the matrix of covariance of the raw data. When features far outnumber instances, the matrix of covariance will take up a lot of memory space, especially for high-dimensional data. Meanwhile, the matrix of covariance will magnify errors in the raw data to some extent. Thirdly, the DPVSD algorithm computes the singular values and singular matrices by SVD, while the existing private PCA algorithms use EVD. Generally, SVD can be considered a black box and has higher execution efficiency compared with EVD, although the two decomposition methods generate the same projection subspace by singular vectors or eigenvectors under the nonprivate situation. There are similar computing processes with the EVD in the next three steps. Lastly, the DPSVD distributes the private classification model to predict the new instances, prior to idea that it projects them to the same singular subspace by the private singular vectors. In brief, the DPSVD trains a private SVM classifier for predicting the new instances in future.

*4.2. Privacy Analysis.* Firstly, the sensitivity of the function  $q(D)$  is analyzed, and then the DPSVD is demonstrated to satisfy  $(\epsilon, \delta)$ -differential privacy. In the DPSVD algorithm, the noise matrix is added to the data matrix  $D$ ; therefore  $q(D) = D$ . Given that two adjacent data matrices  $D$  and  $D'$  differ by exactly one row corresponding to an instance, we set  $D'$  obtained from  $D$  by deleting the last row,  $D = [x_1, \dots, x_n]^T \in R^{n \times d}$  and  $D' = [x_1, \dots, x_{n-1}]^T \in R^{(n-1) \times d}$ , and assume each row has unit  $\ell_2$  norm  $\|x_i\|_2 \leq 1$  at the most.

**Lemma 1.** *The sensitivity of the function  $q(D)$   $S_2$  equals one.*

*Proof.* According to Definition 2, it obtains  $S_2$  by the following inequation:

$$S_2 = \max_{D, D'} \|q(D) - q(D')\|_2 = \max_{D, D'} \|D - D'\|_2 = \max_{D, D'} \|x_n\|_2 \leq 1 \quad (13)$$

Therefore, the sensitivity of the function  $q(D)$  equals one.

<p><b>Input:</b> Raw data matrix <math>D \in R^{n \times d}</math>, instances <math>n</math>, features <math>d</math>, privacy parameters <math>\epsilon</math>, <math>\delta</math> and <math>\beta</math>, accumulative contribution rate of principal components <math>\gamma</math>;</p> <p><b>Output:</b> Classification model <math>f(x) = \sum_{i=1}^n \alpha_i y_i K(x_i, x) + b</math>, private singular vectors <math>V_k</math>;</p> <p><b>Begin</b></p> <p>Generate a noise matrix <math>E \in R^{n \times d}</math>, every entry is i.i.d. and sampled from <math>N(0, \beta^2)</math>;</p> <p>Add the noise matrix to the raw data matrix <math>D' = D + E</math>;</p> <p>Compute the singular values <math>\sigma</math> and singular matrices <math>U, V</math> of <math>D'</math> by SVD, <math>D' = USV^T</math>;</p> <p>Select the target dimension <math>k</math> according to <math>\sum_{i=1}^k \sigma_i^2 / \sum_{i=1}^d \sigma_i^2 \geq \gamma</math>;</p> <p>Select first <math>k</math> singular vectors <math>V_k</math> to project the original training instances to the low-dimensional singular subspace <math>Y = DV_k</math>;</p> <p>Compute the classification model <math>f(x)</math> in the singular subspace;</p> <p>Use <math>f(x)</math> and <math>V_k</math> to predict the new instances.</p> <p><b>End</b></p>
--

ALGORITHM 1: DPSVD.

TABLE 2: The comparison between the three algorithms.

Algorithm	PCA	Adding mode	Noise form	Noise scale	Mechanism	Privacy level
DPSVD	SVD	$D$	Asymmetric	$O(\sqrt{d}/n\epsilon)$	Gaussian	$(\epsilon, \delta)$
AG	EVD	$D^T D/n$	Symmetric	$O(\sqrt{d}/n\epsilon)$	Gaussian	$(\epsilon, \delta)$
DPPCA-SVM	EVD	$D^T D/n$	Symmetric	$O(d/n\epsilon)$	Laplace	$(\epsilon, 0)$

TABLE 3: Test datasets.

Indices	Datasets	Instances	Features	Ranges
1	A1a	1605	119	[0, 1]
2	Mushrooms	8124	112	[-1, 1]
3	Musk	6598	166	[-1, 1]
4	Splice	1000	60	[-1, 1]

**Theorem 1.** *The DPSVD satisfies  $(\epsilon, \delta)$ -differential privacy.*

*Proof.* To demonstrate that the DPSVD satisfies  $(\epsilon, \delta)$ -differential privacy, it is necessary to demonstrate every step in the algorithm satisfies it. According to Lemma 1, it obtains that  $S_2$  equals one. Let  $\beta = S_2 \sqrt{2 \ln(1.25/\delta)}/\epsilon = \sqrt{2 \ln(1.25/\delta)}/\epsilon$ ; then Step (1) and Step (2) satisfy DP according to Definition 4. Step (3) and Step (4) postprocess the private data matrix  $D'$ ; they also satisfy DP. Step (5) generates the private singular vectors  $V_k$ ; the projected instances  $Y$  in the low-dimensional singular subspace are private as well. Meanwhile,  $Y$  does not need to be distributed to users. Step (6) and Step (7) compute the classification model based on private projection instances and distribute it together with private singular vector to predict the new instances. The last three steps do not violate the privacy requirement of DP. Therefore, the DPSVD satisfies  $(\epsilon, \delta)$ -differential privacy.

**4.3. Algorithm Comparison.** The three algorithms were compared theoretically between DPSVD, AG [9], and DPPCA-SVM [26] summarized in Table 2. Other ones have been compared by the DPPCA-SVM algorithm. Our algorithm uses SVD to perform PCA; it does not need to

compute matrix of covariance and symmetrize the noise matrix as the description above. It obtains the same noise scale as AG algorithm, because they use identical mechanism of DP to generate the noise matrix.

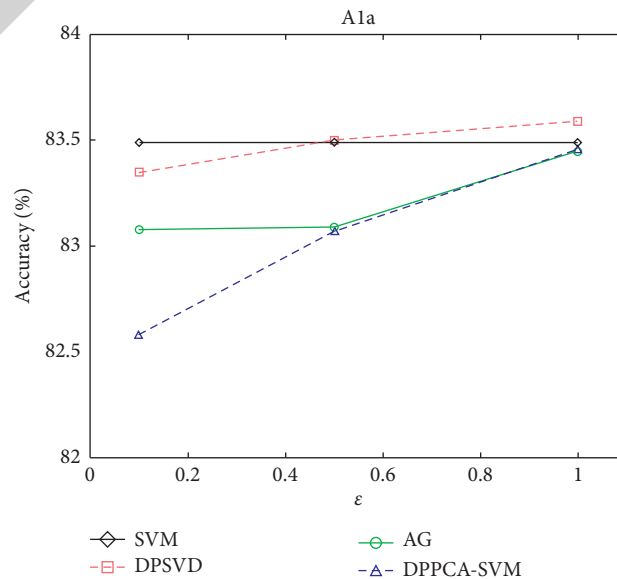
Therefore, the classification model and the singular vectors for projection are both private; they can be used to predict the new instances in the same singular subspace. The main advantage of the DPSVD compared with other private SVMs is that our algorithm trains the classification model in the private low-dimensional singular subspace generated by SVD. In this way, the features of the instances in the singular subspace become linearly independent and low-dimensional and therefore have higher time and space efficiency for training the classification model. The difference between our algorithm with other private PCA algorithms is that it does not need to calculate the matrix of covariance or symmetrize the noise matrix. Meanwhile, the DPSVD protects privacy of the training instances before training the classification model; many optimization methods of SVMs can be applied directly to the training progress.

## 5. Results

**5.1. Datasets.** Our experiments were carried out on four popular datasets for testing SVM performance. Table 3

TABLE 4: Performance comparison of algorithms on different datasets.

Datasets	$\epsilon$	Algorithm	Accuracy				SV			
			Mean	Std	Max	Min	Mean	Std	Max	Min
A1a	--	SVM	83.49	--	--	--	754	--	--	--
	0.1	DPSVD	<b>83.35</b>	0.15	83.61	83.24	<b>763</b>	6	771	756
		AG	83.08	0.24	83.30	82.68	743	4	747	738
		DPPCA-SVM	82.58	0.58	83.30	81.93	771	12	783	754
	0.5	DPSVD	<b>83.50</b>	0.22	83.86	83.30	<b>755</b>	9	768	743
		AG	83.09	0.16	83.18	82.80	686	3	689	682
		DPPCA-SVM	83.07	0.50	83.49	82.24	764	14	778	747
	1	DPSVD	<b>83.59</b>	0.13	83.74	83.43	<b>753</b>	5	759	748
		AG	83.45	0.24	83.80	83.18	697	4	702	693
DPPCA-SVM		83.46	0.41	84.11	82.99	756	14	778	742	
Mushrooms	--	SVM	99.90	--	--	--	617	--	--	--
	0.1	DPSVD	<b>99.89</b>	0.01	99.90	99.88	<b>639</b>	39	700	607
		AG	99.18	0.02	99.21	99.15	518	3	521	514
		DPPCA-SVM	99.49	0.42	99.89	99.02	683	67	747	604
	0.5	DPSVD	<b>99.90</b>	0.01	99.90	99.89	<b>633</b>	25	674	607
		AG	99.19	0.05	99.26	99.14	524	5	531	517
		DPPCA-SVM	99.59	0.38	99.90	99.06	763	50	811	687
	1	DPSVD	<b>99.90</b>	0.00	99.90	99.90	<b>602</b>	26	625	559
		AG	99.79	0.04	99.83	99.73	445	22	469	417
DPPCA-SVM		99.83	0.08	99.98	99.78	651	81	779	559	
Musk	--	SVM	93.95	--	--	--	1351	--	--	--
	0.1	DPSVD	<b>94.08</b>	0.11	94.23	93.95	<b>1351</b>	15	1369	1330
		AG	88.96	0.08	89.09	88.89	1865	10	1876	1855
		DPPCA-SVM	93.97	0.20	94.23	93.74	1379	8	1391	1369
	0.5	DPSVD	<b>94.14</b>	0.11	94.27	94.00	<b>1359</b>	14	1379	1341
		AG	88.94	0.01	88.95	88.92	1866	6	1874	1858
		DPPCA-SVM	94.10	0.15	94.35	93.97	1336	15	1355	1315
	1	DPSVD	94.14	0.10	94.29	94.04	<b>1345</b>	10	1358	1333
		AG	88.93	0.02	88.95	88.91	1872	10	1887	1860
DPPCA-SVM		<b>94.19</b>	0.17	94.35	93.92	1318	40	1384	1275	
Splice	--	SVM	94.30	--	--	--	607	--	--	--
	0.1	DPSVD	<b>91.08</b>	0.75	92.40	90.60	635	17	662	616
		AG	90.56	0.83	91.30	89.30	<b>591</b>	16	605	568
		DPPCA-SVM	87.14	0.32	87.40	86.70	643	16	660	619
	0.5	DPSVD	92.00	0.80	92.80	90.70	<b>610</b>	10	625	600
		AG	<b>93.58</b>	0.38	94.00	93.00	588	5	595	582
		DPPCA-SVM	87.22	0.73	88.40	86.60	659	34	706	615
	1	DPSVD	92.36	0.56	93.10	91.80	<b>618</b>	19	645	594
		AG	<b>93.56</b>	0.31	93.90	93.10	594	3	599	591
DPPCA-SVM		87.36	0.94	88.30	86.00	641	21	667	614	

FIGURE 1: Accuracy at various  $\epsilon$  on dataset A1a.



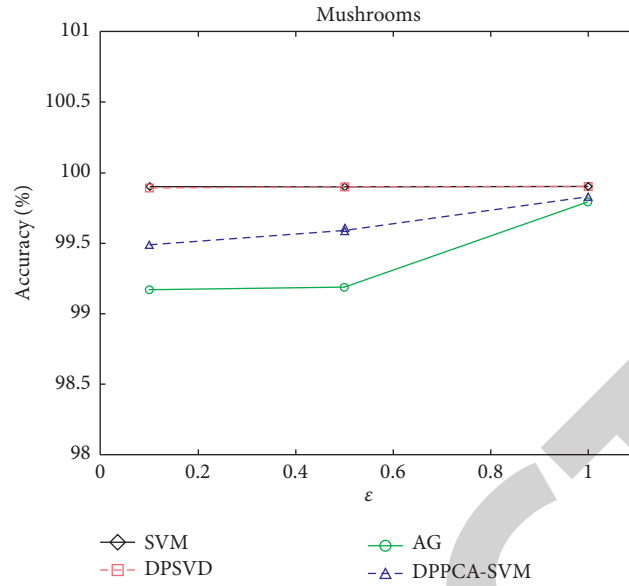


FIGURE 2: Accuracy at various  $\epsilon$  on dataset Mushrooms.

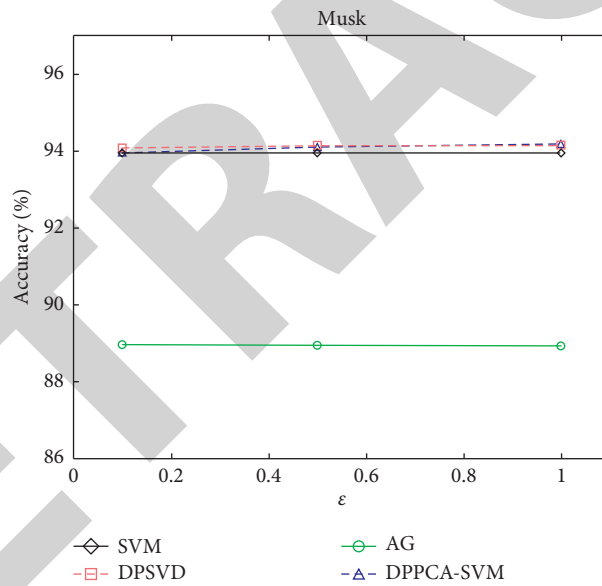


FIGURE 3: Accuracy at various  $\epsilon$  on dataset Musk.

describes their basic information, including the number of the instances, the number of the features, and the ranges of values in the data. They are accessible at <https://www.csie.ntu.edu.tw/~cjlin/libsvmtools/datasets/> and <http://archive.ics.uci.edu/ml/datasets.php>. It trains the SVM to compare the performance of algorithms based on LIBSVM (version

3.25) [36] with the radial basis function as the kernel function and default parameters.

5.2. Algorithm Performance Experiments. The performance of the DPSVD was compared with AG, DPPCA-SVM, and the nonprivate SVM on the four real datasets. In the

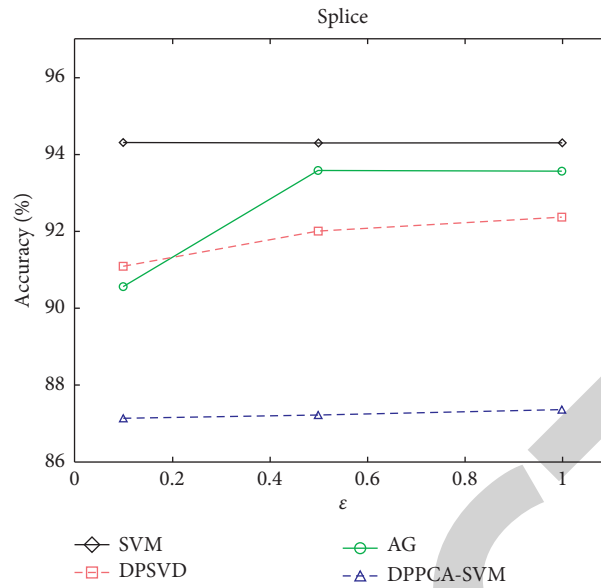


FIGURE 4: Accuracy at various  $\epsilon$  on dataset Splice.

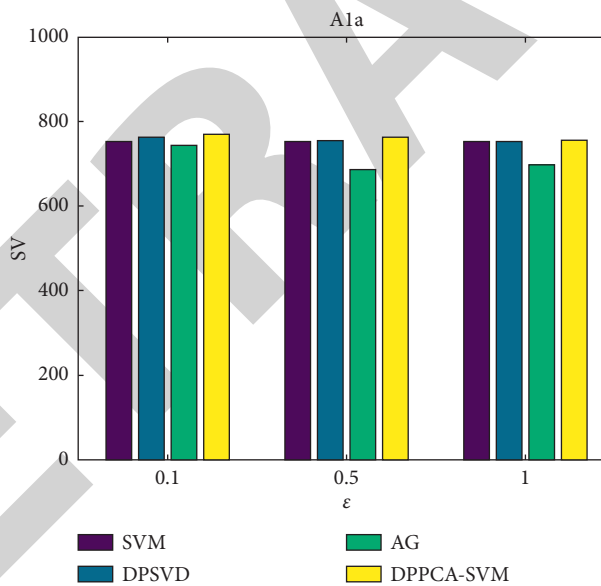


FIGURE 5: SV at various  $\epsilon$  on dataset A1a.

experiments, it designed two metrics of algorithm performance *Accuracy* and *SV*. *Accuracy* denotes how accurate the classification is, and *SV* denotes how many SVs are contained in the classifier. The higher the *Accuracy*, the greater the usability of the classifier. The closer *SV* to that of nonprivate SVM, the better the stability of the algorithm. The privacy budget  $\epsilon$  was set at 0.1, 0.5, and 1,  $\delta$  at  $1/n^2$ , and the accumulative contribution rate of principal components

$\gamma$  at 90%. Three private algorithms were implemented five times under every privacy budget parameter. The mean values, standard deviation, and maximum and minimum values of the two metrics are given in Table 4.

From the experiments results in Table 4, the DPSVD was the most accurate in classification than the other two private classifiers under different privacy budget for most of the datasets. Sometimes, our algorithm even outperformed the

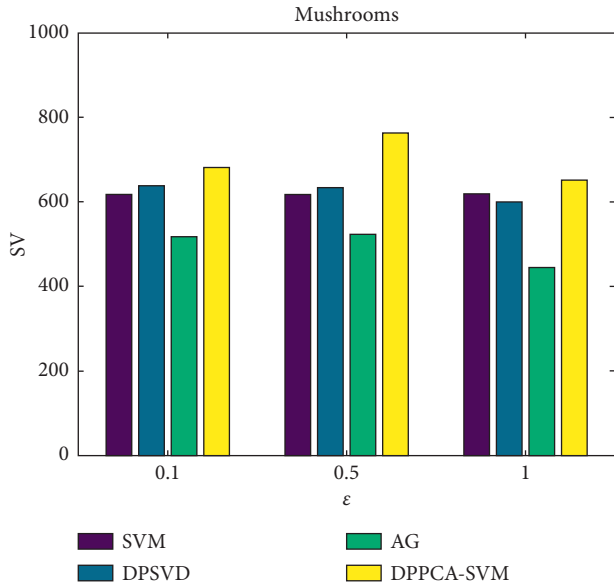


FIGURE 6: SV at various  $\epsilon$  on dataset Mushrooms.

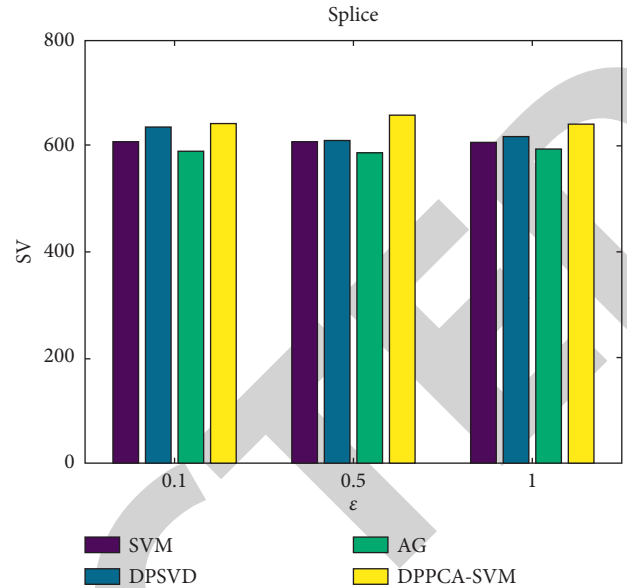


FIGURE 8: SV at various  $\epsilon$  on dataset Splice.

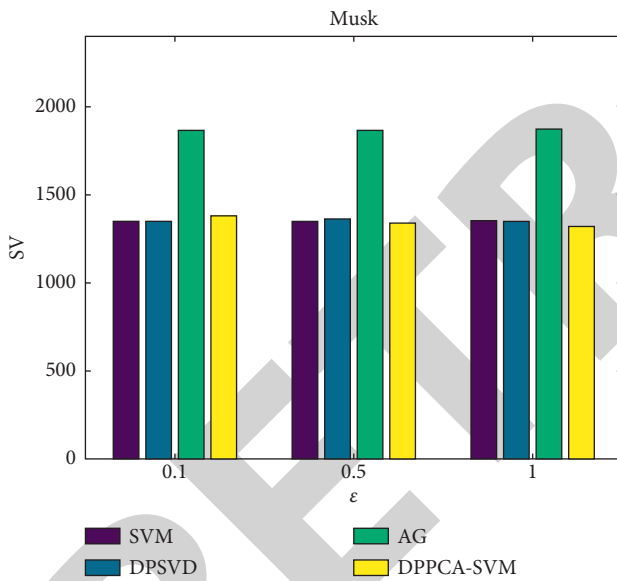


FIGURE 7: SV at various  $\epsilon$  on dataset Musk.

nonprivate SVM. This is mainly because our algorithm removes the linear dependence between features and unimportant features by SVD. Meanwhile, our algorithm has the better stability of the algorithm as its SV is much closer to the nonprivate SVM. To compare algorithm performance more intuitively, the mean values of the two metrics for the four algorithms are shown in Figures 1–8.

In Figure 1 to Figure 3, the DPSVD achieved the highest classification accuracy in the three private algorithms and closer to the nonprivate SVM than the other two algorithms. In Figure 4, the AG achieved the higher classification accuracy than the DPSVD as the privacy budget increases. In Figure 5 to Figure 8, the DPSVD contained the closer

number of SVs in the classifier to the nonprivate SVM than the other two algorithms. Therefore, the PDSVD achieved higher classification accuracy and better algorithm stability on most of the datasets and approximated the performance of the nonprivate SVM. The AG algorithm on dataset Musk in Figure 3 and the DPPCA-SVM algorithm on dataset Splice in Figure 4 have relatively lower classification accuracy. It also shows that the DPSVD has better algorithm stability.

## 6. Conclusions

To solve the privacy leak of SVM classifiers, especially on high-dimensional data, the DPSVD algorithm was proposed to project the training instances into the low-dimensional singular subspace and train a private SVM classifier on it while not violating the privacy requirements for the training data. The DPSVD is proved to satisfy DP. The main advantages of the DPSVD include three aspects. Firstly, it trains the classification model in the private low-dimensional singular subspace; therefore it has higher time and space efficiency compared with other private SVMs. Secondly, it does not need to calculate the matrix of covariance or symmetrize the noise matrix and has higher classification accuracy and better stability of the algorithm than other existing private PCA algorithms through the comparison experiments. Thirdly, it protects privacy of the training instances before training the classification model, and many optimization methods of SVMs can be applied directly to the training progress. Meanwhile, its algorithmic ideas can be applied to other machine learning areas to solve data privacy problems. However, the DPSVD can only solve the linear dependence between the data features. In future work, we will consider the nonlinear dependence to train a private classification model. In addition, the problem of data instances compression through SVD is another research direction.

## Data Availability

The raw data of the four datasets are available at <https://www.csie.ntu.edu.tw/~cjlin/libsvmtools/datasets/> and <http://archive.ics.uci.edu/ml/datasets.php>.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grants 61672179, 61370083, 61402126, and 61501275, by the Natural Science Foundation of Heilongjiang Province under Grant F2015030, by the Science Fund for Youths of Heilongjiang Province under Grant QC2016083, by the Postdoctoral Fellowship of Heilongjiang Province under Grant LBH-Z14071, and by the Fundamental Research Funds in Heilongjiang Provincial Universities under Grant 135509312.

## References

- [1] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining and Knowledge Discovery*, vol. 2, no. 2, pp. 121–167, 1998.
- [2] K.-P. Lin and M.-S. Chen, "On the design and analysis of the privacy-preserving SVM classifier," *IEEE Transactions on Knowledge and Data Engineering*, vol. 23, no. 11, pp. 1704–1717, 2011.
- [3] C. Dwork, "Differential privacy," *Automata, Languages and Programming*, in *Proceedings of the 33th Colloquium on Automata, Languages and Programming*, pp. 1–12, Venice, Italy, July 2006.
- [4] C. Liu, J. Yang, W. Zhao et al., "Differential privacy protection of face images based on region growing," *Traitement du Signal*, vol. 38, no. 5, pp. 1385–1401, 2021.
- [5] J. Yu, "Design of a privacy-preserving algorithm for peer-to-peer network based on differential privacy," *Ingénierie des Systèmes d'Information*, vol. 24, no. 4, pp. 433–437, 2019.
- [6] C. Liu, J. Yang, Y. Zhang et al., "Non-global privacy protection facing sensitive areas in face images," *Traitement du Signal*, vol. 38, no. 6, pp. 1677–1687, 2021.
- [7] T. Zhu, G. Li, W. Zhou, and P. S. Yu, "Differentially private data publishing and analysis: a survey," *IEEE Transactions on Knowledge and Data Engineering*, vol. 29, no. 8, pp. 1619–1638, 2017.
- [8] C. Dwork, F. McSherry, K. Nissim, and A. Smith, "Calibrating noise to sensitivity in private data analysis," *Theory of Cryptography*, in *Proceedings of the 3rd Conference on theory of Cryptography*, pp. 265–284, New York, NY, USA, 4–7 March 2006.
- [9] C. Dwork, K. Talwar, A. Takurta, and L. Zhang, "Analyze gauss: optimal bounds for privacy-preserving principal component analysis," in *Proceedings of the Forty-Sixth Annual ACM Symposium on Theory of Computing*, pp. 11–20, ACM, New York, NY, USA, May 31 - June 3, 2014.
- [10] F. McSherry and K. Talwar, "Mechanism design via differential privacy," in *Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science (FOCS'07)*, Providence, RI, USA, 21–23 Oct. 2007.
- [11] D. Kalman, "A singularly valuable decomposition: the SVD of a matrix," *The College Mathematics Journal*, vol. 27, no. 1, pp. 2–23, 1996.
- [12] K. Chaudhuri and C. Monteleoni, "Privacy-preserving logistic regression," in *Proceedings of the 2008 21st ACM Conference on Advances in Neural Information Processing Systems*, pp. 289–296, Austin, TX, USA, November 2009.
- [13] K. Chaudhuri, C. Monteleoni, and A. D. Sarwate, "Differentially private empirical risk minimization," *Journal of Machine Learning Research : JMLR*, vol. 12, no. 3, pp. 1069–1109, 2011.
- [14] B. Ip Rubinstein, P. L. Bartlett, L. Huang, and N. Taft, "Learning in a large function space: privacy-preserving mechanisms for SVM learning," 2009, <https://arxiv.org/abs/0911.5708>.
- [15] H. Li, L. Xiong, L. M. Ohno, and X. Jiang, "Privacy preserving RBF kernel support vector machine," *BioMed Research International*, vol. 2014, p. 10, Article ID 827371, 2014.
- [16] Y. L. Zhang, Z. F. Hao, and S. P. Wang, "A differential privacy support vector machine classifier based on dual variable perturbation," *IEEE Access*, vol. 7, p. 98251, 2019.
- [17] X. Liu, Q. Li, and T. Li, "Private classification with limited labeled data," *Knowledge-Based Systems*, vol. 133, no. 1, pp. 197–207, 2017.
- [18] Z. Sun, J. Yang, X. Li, and J. Zhang, "DPWSS: differentially private working set selection for training support vector machines," *PeerJ Computer Science*, vol. 7, p. e799, 2021.
- [19] Z. Sun, J. Yang, X. Li, and J. Zhang, "Differentially private kernel support vector machines based on the exponential and Laplace hybrid mechanism," *Security and Communication Networks*, vol. 2021, p. 19, Article ID 9506907, 2021.
- [20] A. Blum, C. Dwork, F. McSherry, and K. Nissim, "Practical privacy: the SuLQ framework," in *Proceedings of the twenty-fourth ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*, pp. 128–138, ACM, New York, NY, USA, June 2005.
- [21] K. Chaudhuri, A. D. Sarwate, and S. Kaushik, "A near-optimal algorithm for differentially-private principal components," *Journal of Machine Learning Research*, vol. 14, no. 1, pp. 2905–2943, 2005.
- [22] H. Imtiaz and A. D. Sarwate, "Symmetric matrix perturbation for differentially-private principal component analysis," in *Proceedings of the 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2339–2343, ICASSP, Shanghai, China, March 2016.
- [23] H. Imtiaz and A. D. Sarwate, "Differentially private distributed principal component analysis," in *Proceedings of the 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2206–2210, 15–20 April 2018.
- [24] W. Jiang, C. Xie, and Z. Zhang, "Wishart mechanism for differentially private principal components analysis," in *Proceedings of the Birtith AAI Conference on Artificial Intelligence*, AAAI Press, Phoenix, AZ, USA, February 2016.
- [25] Y. H. Xu, G. Yang, and S. J. Bai, "Laplace input and output perturbation for differentially private principal components analysis," *Security and Communication Networks*, vol. 2019, p. 10, Article ID 9169802, 2019.
- [26] Y. X. Huang, G. Yang, Y. H. Xu, and H. Zhou, "Differential privacy principal component analysis for support vector machines," *Security and Communication Networks*, vol. 2021, p. 12, Article ID 5542283, 2021.
- [27] T. Huang, R. Zhao, L. Bi, D. Zhang, and C. Lu, "Neural embedding singular value decomposition for collaborative filtering," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–9, 2021.

## Research Article

# Application of a Hybrid Model of Big Data and BP Network on Fault Diagnosis Strategy for Microgrid

Chundi Jiang<sup>1,2</sup> and Zhiliang Xia<sup>3</sup> 

<sup>1</sup>Logistics Engineering College, Shanghai Maritime University, Shanghai 201306, China

<sup>2</sup>Electrical and Information Engineering, Quzhou University, Quzhou 324000, China

<sup>3</sup>Wenzhou Polytechnic, Wenzhou 325035, China

Correspondence should be addressed to Zhiliang Xia; 2011041047@wzpt.edu.cn

Received 22 January 2022; Revised 21 February 2022; Accepted 22 February 2022; Published 25 March 2022

Academic Editor: Daqing Gong

Copyright © 2022 Chundi Jiang and Zhiliang Xia. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Aiming at the characteristics of timely transmission, rapid update, and large magnitude of microgrid data, based on the large data samples generated by microgrid operation, a fault diagnosis and analysis method of microgrid systems supported by big data is proposed in this paper. The multisource joint feature vectors of microgrid are extracted using Wavelet transform, Rayleigh entropy, and big data technology, which combine short-circuit current and voltage. The extracted feature dataset is clustered and segmented to realize deep data mining. Combining BP neural network and big data, the fault diagnosis of microgrid is realized. The simulation results show that the BP neural network algorithm based on big data support can accurately identify the type and phase of internal faults in microgrid, which is more suitable for extracting the temporal characteristics of information and spatiotemporal correlation of data to realize the prediction of big data and solve the core problems in the analysis of big data of microgrid faults, and the accuracy is as high as 96.8%.

## 1. Introduction

Big data refers to the dataset that cannot be captured, managed, and processed by conventional software tools within a certain period of time. It is a massive and diversified information asset of high growth rate that needs a new processing mode to have stronger decision-making power, insight, and process optimization ability. “Volume, Variety, Value, Velocity, and Veracity” are the “5V” characteristics of big data. Relevant research and surveys have pointed out that the global annual data growth rate is basically twice or even higher than that of the previous year. In the next 10 years, nonstructural data will account for about 90% and the data patterns will be different. It will become impossible to analyze based on previous experience. Therefore, it is necessary to study relevant data mining technologies and understand and master the basic “5V” characteristics of big data, which is particularly important for data analysis based on big data. BP neural

network can be applied to the prediction of models and the study of the relationship between different models. Therefore, the combination of big data technology and BP neural network can deal with large and complex nonlinear structural problems from a statistical point of view, with high stability and accuracy. Big data fault analysis refers to collecting a large amount of data through real-time analysis and mining of microgrid fault information to master the microgrid operation characteristics, accurately predict the microgrid topology behavior, and improve the capabilities of service and risk control. The key to big data is to be able to quickly obtain useful information from a large amount of data or quickly realize big data assets. Therefore, the information processing of big data is often based on cloud computing. Cloud computing is the product of a new era. Taking cloud computing as the development strategy, big data is applied to cloud computing to obtain a series of data that can be used for fault analysis, so as to form a big data fault diagnosis model.

The reasonable development and utilization of green and clean energy have become an important topic with increasingly prominent energy crisis, environmental pollution, and slow recovery caused by large power grids. Microgrid has been widely used because of its flexible installation locations, less pollution, and high energy efficiency. Its technical problems have also attracted the extensive attention of researchers at home and abroad. One of the important research fields is fault diagnosis technology, which guarantees safe and reliable operation. The common fault types of internal lines in the microgrid are single-phase grounding short circuit, two-phase short circuit, three-phase short circuit, etc. Conventional fault diagnosis methods cannot be directly applied to microgrid and work well due to the difference between current and voltage and bidirectional power flow. It is an urgent problem to find new fault diagnosis and identification methods for microgrid because of the changing characteristics of fault voltage and current.

A lot of research has been carried out and been applied in practice to the fault diagnosis technology of microgrid. The research mainly focuses on the following two aspects: Firstly, fault diagnosis is carried out according to the changes of circuit breakers, the flexible topology, protection elements, and other equipment status in the microgrid. For example, the voltage and bus current signals of the low-voltage circuit breaker are used for fault diagnosis. Using SOM neural network and multiagent systems, the simulation results show that good diagnosis results can be achieved for a single fault, but the situation of multiple faults in the same period needs further study [1, 2]. In [3], a Petri net analysis model using line fault protection information is established. The corresponding protection set information needs to be updated while the topology changes without remodeling, which can adapt to the variability of the microgrid topology structure. However, the identification accuracy of fault diagnosis will be affected when the protection device and circuit breaker have such factors as refusing to operate or misoperation.

Secondly, the abnormal change of voltage or current is used to diagnose the fault in the microgrid. The transient recovery performances of microgrid under different control modes are compared and analyzed by studying the change of current amplitude and the transient duration of overvoltage when a short-circuit fault occurs in the microgrid [4, 5]. According to the impedance and impedance angle, the symmetrical component method was used to analyze the voltage and current in the microgrid during a fault. However, it is difficult to accurately extract the fault component when the system is unstable, is oscillating, has interference, or has continuous fault. At the same time, its effect is a lack of experimental verification while only theoretical and methodological research is carried out.

The fault voltage and current signal are used to realize fault diagnosis and identification in a combined wavelet packet with a neural network, but the fault phase cannot be accurate to the line [6, 7]. The voltage signal of the optical storage microgrid line is analyzed to realize the safe and reliable operation of the optical storage microgrid based on genetic algorithm [8]. Another fault

detection scheme based on deep neural networks and wavelet transform for microgrid was proposed in [9]. The authors in [10] employed an approach focusing on identifying and evaluating the faulted line section by implementing data mining and wavelet packet transform. At present, the scale of microgrid data has increased exponentially, and the traditional data processing methods cannot afford to process large-scale data. At the same time, the big data of the microgrid system runs through all nodes of smart grid, which is far beyond the scope of traditional power system monitoring. The current data processing platform is difficult to meet the requirements of smart grid for the processing of power system data. In particular, it is difficult to achieve real-time data, which results in the loss and repetition of power grid data. In the era of big data, how to process microgrid data to analyze and diagnose the working conditions of power equipment is an urgent task for the development of intelligent microgrids. In order to solve the problems of incomplete fault information, single information structure, and imperfect diagnosis results in the existing microgrid fault diagnosis methods, this paper extracts and reconstructs the features of multisource data using BP neural network based on big data technology, obtains the spatiotemporal correlation characteristics between different data, and then realizes fault prediction and diagnosis better.

The main contributions of this paper are as follows: First, most of the existing research literature is about the relationship between the amplitude of voltage, current, and fault, but less about the relationship between the time-frequency characteristics of voltage and current and the impact of time-frequency characteristics on fault. In fact, the time-frequency characteristics of voltage and current on the fault will have a greater impact. Secondly, this paper not only analyzes the relationship between voltage and current but also studies the regulatory role of voltage, current, and fault. Third, this paper combines BP neural network and big data to predict the spatiotemporal correlation of data in order to realize fault diagnosis using big data analysis. Considering that the voltage and current of three-phase lines contain rich transient sudden change signals when the internal lines of the microgrid fail, which can effectively reflect the fault characteristics, this paper obtained the high-frequency and low-frequency details of the signal using wavelet analysis, Rayleigh entropy, and other theoretical methods to extract the characteristic vectors of three-phase line fault voltage and current from massive microgrid data. Based on the analysis of fault information, the BP neural network is applied to fault diagnosis and identification. Trained with historical data as samples, the BP neural network model supported by big data is constructed to realize the accurate diagnosis of fault type and phase.

The rest of this paper is arranged as follows: Section 2 describes fault characteristics; Section 3 gives wavelet packet decomposition and energy entropy; Section 4 provides construction of the BP neural network fault diagnosis model; Section 5 describes fault diagnosis algorithm and process;



Section 6 gives simulation analysis and experimental verification; Section 7 concludes the study.

## 2. Fault Characteristics Analysis

**2.1. Microgrid Prototype.** A microgrid prototype is shown in Figure 1. The bus voltage is 220 V and connected to the distribution network through the PCC of the common connection point. DG1 is a photovoltaic power generation unit with a capacity of 60 kVA. DG2 is a wind power source with a capacity of 20 kVA. There are three power transmission lines in the microgrid: WL1 is 1 km long with 8 kW load; WL2 is 5 km with 50 kW load; and WL3 is 2 km long with 15 kW load. The fault resistance and grounding resistance are both set as 0.001  $\Omega$  when single-phase faults happen, while the grounding resistance is set as 10 m $\Omega$  when two-phase or three-phase short circuits occur.

The three-phase output voltage and current are symmetrical when the microgrid operates normally. The current waveform is shown in Figure 2. The three-phase output current is equal, and the phase difference is 120°.

**2.2. Fault Analysis.** The Matlab simulation platform is used to analyze the variation of line voltage and current under different fault types and fault positions. In this paper, Matlab/Simulink 2018b software is mainly used in the process of modeling and simulation. Simulink is mainly responsible for the construction of the whole system and offline simulation verification. At the same time, it can be combined with StarSim HIL and StarSim RCP software of Yuankuan; then, the hardware in the loop test is realized. The simulation analysis is carried out at 10%, 50%, 70%, and 85% of the microgrid side, respectively, while single-phase ground fault, phase-to-phase short circuit, two-phase ground fault, or three-phase short-circuit fault happens. The simulation time is set as 0.5 s, and the fault will be removed in 0.4 s after the fault occurs in 0.1 s. When the fault occurs, the voltage and current at the fault point are shown in Figure 3.

Figure 3 shows that the voltage at the fault point decreases significantly after the short circuit, the amplitude of fault current increases, and there are certain components of harmonic and aperiodic components. It is very likely to cause microgrid instability and protection misoperation even after the fault is removed. The fault diagnosis method based on the single signal has certain limitations, so it is necessary to consider the voltage and current signals to build a new feature vector for fault diagnosis.

The fault data information during microgrid operation includes steady-state data, parameter data, alarm events, etc. The circuit information includes frequency, voltage, current, harmonic voltage, harmonic current, voltage imbalance, current imbalance, flicker, power and power factor in the circuit, power grid clutter interference, vibration, temperature and humidity, harmonic interference, abnormal events, and other indicators. Through the analysis and

processing of these data, accurate fault characteristics can be extracted to realize fault identification and diagnosis.

## 3. Wavelet Packet Decomposition and Energy Entropy

**3.1. Wavelet Packet Decomposition.** The wavelet packet selects the optimal basis to decompose the original signal in the frequency domain, which improves the ability of signal analysis, avoids the defect of fixed time frequency of wavelet decomposition, and accurately reflects the nature and characteristics of the signal. It has good time-domain and frequency-domain positioning characteristics and excellent signal adaptability.

Let  $\phi(x)$  be an orthogonal scaling function and  $\psi(x)$  be a wavelet function; then, the two-scale function equation is as follows:

$$\begin{aligned}\phi(x) &= \sqrt{2} \sum_k h_k \phi(2x - k), \\ \psi(x) &= \sqrt{2} \sum_k g_k \phi(2x - k),\end{aligned}\quad (1)$$

where  $h_k$  is the scale coefficient and  $g_k$  is the wavelet coefficient.

Defining a basis function  $u_0(x) = \phi(x)$  and  $u_1(x) = \psi(x)$ , then the two-scale equation is generalized as follows:

$$\begin{cases} u_{2n}(t) = \sqrt{2} \sum_{k \in \mathbb{Z}} h_k u_n(2t - k), \\ u_{2n+1}(t) = \sqrt{2} \sum_{k \in \mathbb{Z}} g_k u_n(2t - k), \end{cases}\quad (2)$$

where the constructed sequence  $\{u_n(x)\}$  is the wavelet packet of basis function  $u_0(x) = \phi(x)$ .

The parameter  $j$  is the scale index,  $k$  is the position index, and  $n$  is the oscillation number; then,

$$\begin{cases} d_k^{j,2n} = \sum_{i \in \mathbb{Z}} h_{k-2i} d_i^{j+1,n}, \\ d_k^{j,2n+1} = \sum_{i \in \mathbb{Z}} g_{k-2i} d_i^{j+1,n}. \end{cases}\quad (3)$$

The wavelet packet reconstruction algorithm is as follows:

$$d_k^{j+1,n} = \sum_{i \in \mathbb{Z}} (h_{k-2i} d_i^{j,2n} + g_{k-2i} d_i^{j,2n+1}).\quad (4)$$

Taking 3-level wavelet packet decomposition as an example, the process of wavelet packet decomposition and reconstruction is shown in Figure 4. Node  $(i, j)$  represents the  $j$ th node in layer  $i$  ( $i = 0, 1, 2, 3; j = 0, 1, 2, \dots, 7$ ), and each node represents a signal with certain characteristics. For example, the node  $(0, 0)$  represents the original signal  $S$ , node  $(1, 0)$  represents the low-frequency coefficient of the first layer of wavelet packet decomposition, and node  $(1, 1)$  represents the high-frequency coefficient of the first layer.

The relationship between the original signal  $S$  and its decomposition coefficient is as follows:

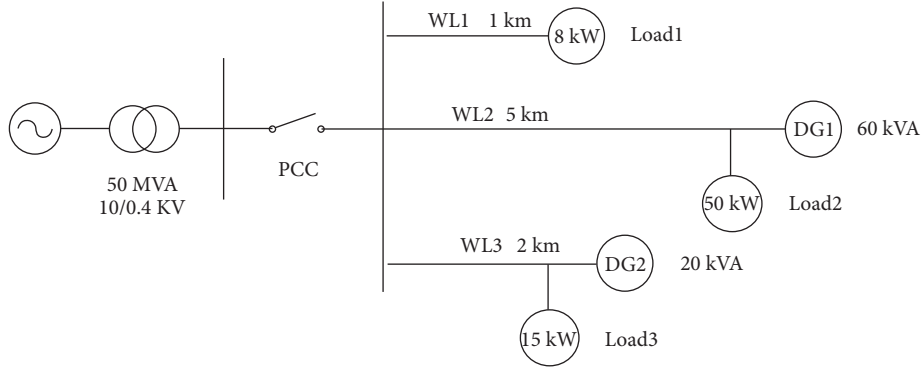


FIGURE 1: System structure.

$$S = S(3, 0) + S(3, 1) + \dots + S(3, 7). \quad (5)$$

**3.2. Feature Entropy Extraction and Multisource Joint Eigenvector Construction.** Entropy can measure the degree of information uncertainty in information theory, such as information entropy and relative entropy [11, 12]. The Shannon entropy extracted from the wavelet packet decomposition and reconstruction can reflect the high-frequency and low-frequency characteristics of the signal more accurately and has stronger anti-interference performance [6]. The time complexity of the frequency component of the signal transient can be accurately expressed by the singular entropy of the Rayleigh wavelet packet, which is more conducive to the identification and diagnosis of fault signals.

Let the random signal  $X = \{x_0, x_1, \dots, x_{N-1}, x_N\}$ ; the probability of occurrence of  $x_i$  is

$$p_i = p(x_i), \quad i = 0, 1, 2, \dots, n-1, \quad (6)$$

$$\sum_{i=0}^N p_i = 1.$$

Then, the Rayleigh entropy of  $X$  is

$$H_q(X) = \frac{1}{1-q} \ln \sum_{i=1}^N p_i^q. \quad (7)$$

The expression of energy entropy fused with wavelet packet decomposition is

$$E_{i,j} = \frac{1}{1-q} \ln \sum_j p_{i,j}^q. \quad (8)$$

The vector  $T$  is composed of the energy entropy of each frequency band:

$$T = [E_{i,0}, E_{i,1}, E_{i,2}, \dots, E_{i,2^n-1}]. \quad (9)$$

To facilitate signal analysis, the vector  $T$  can be normalized and represented by  $T'$  when the energy entropy of each frequency band is relatively large.

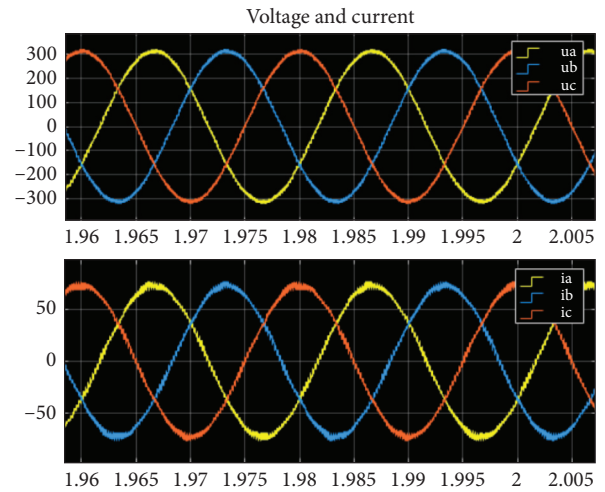


FIGURE 2: Three-phase output current waveform in normal operation.

$$E = \sum_{j=0}^{2^n-1} E_{i,j}, \quad (10)$$

$$T' = \frac{T}{E} = \left[ \frac{E_{i,0}}{E}, \frac{E_{i,1}}{E}, \frac{E_{i,2}}{E}, \dots, \frac{E_{i,2^n-1}}{E} \right].$$

After feature entropy extraction, the voltage and current signals are fused by the interval crossing method to form a new multisource fault feature vector for fault diagnosis. The final signal eigenvector is obtained.

#### 4. Construction of the BP Neural Network Fault Diagnosis Model

In 1986, the BP neural network was proposed by Rumelhart and McClelland, which has become one of the most popular neural network models because of its strong functions of self-learning and adaptability. The BP network can solve the fault diagnosis problem of some complex systems and provide theoretical research and technical implementation methods for more intelligent diagnosis methods.

With the support of big data technology, the development of the BP neural network gradually tends to mature.

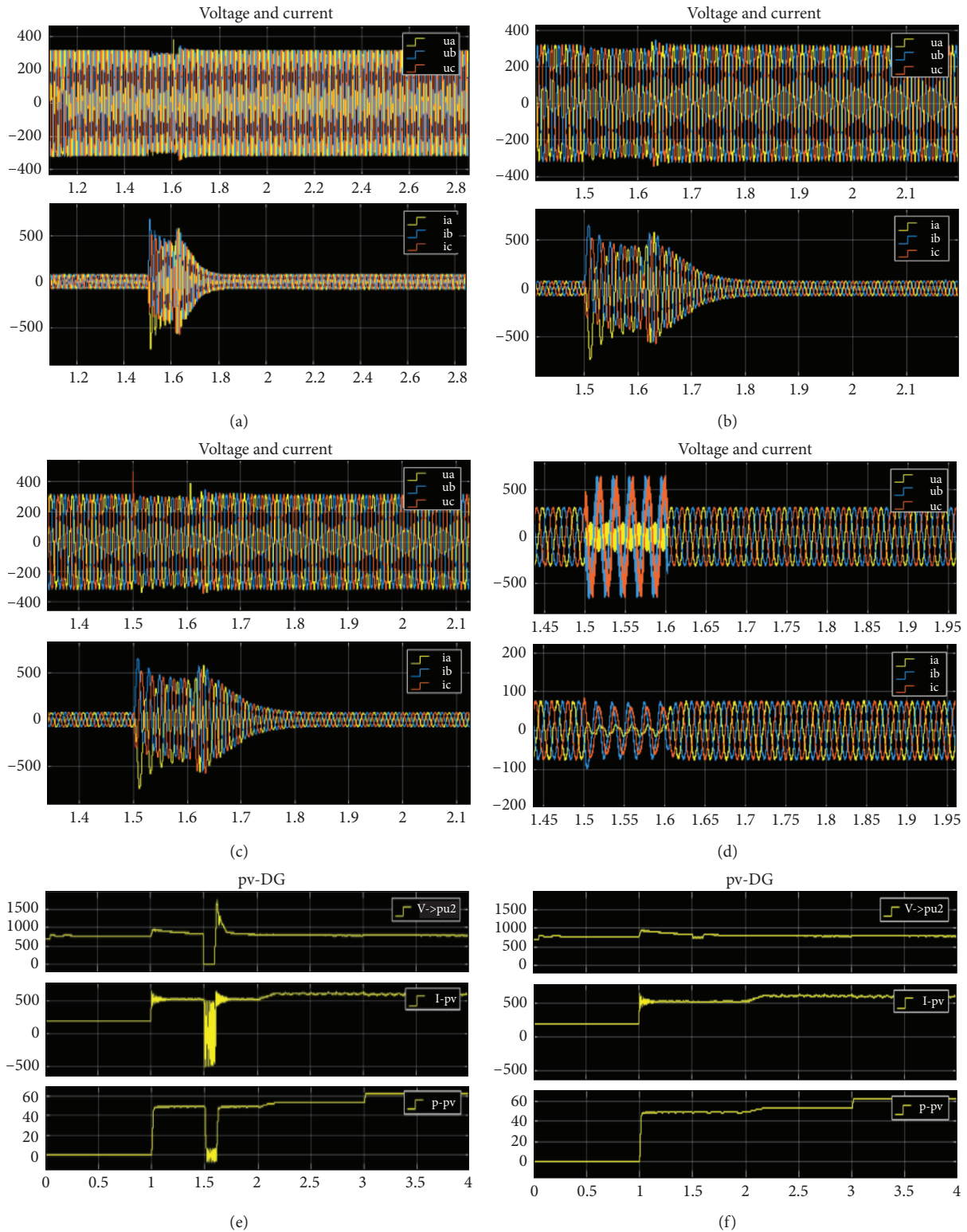


FIGURE 3: Partial fault voltage and current curves: (a) three-phase short circuit (1.5 s–1.6 s); (b) two-phase short circuit (1.5 s–1.6 s); (c) two-phase ground fault (1.5 s–1.6 s); (d) single-phase ground (1.5 s–1.6 s); (e) PV parameters of three-phase fault; (f) PV parameters of single-phase fault.

When analyzing and diagnosing power grid faults, it can convert relay protection information into model input and take the possible faults of microgrid as an output. On this

basis, it completes the selection of the sample set. Then, it cleans and segments the sample, compares the difference between the actual output and the expected output, realizes

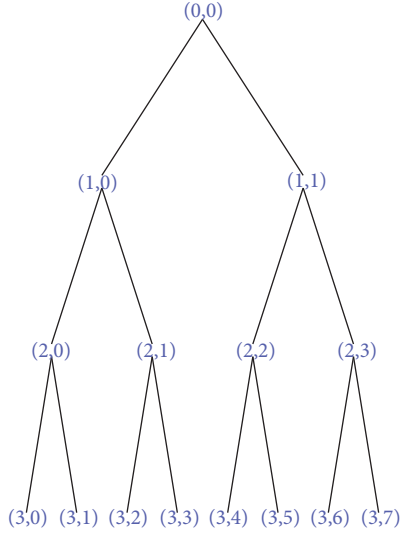


FIGURE 4: 3-level wavelet packet decomposition structure.

the layer-by-layer adjustment of the weight and threshold, and stops after repeated times, and the difference is consistent with the standard. The BP neural network is mainly composed of an input layer, a hidden layer, and an output layer. The topological structure of a typical three-layer BP neural network is shown in Figure 5.

Regulation:  $J$  is the number of nodes in the input layer, the node serial number is  $j$ , input vector  $\dot{X} = [x_1, x_2 \dots x_j \dots x_J]$ ,  $j = 1, 2 \dots J$ ,  $K$  is the number of nodes in the output layer, node serial number is  $k$ ,  $\dot{O} = [o_1, o_2 \dots o_k \dots o_K]$  is the output vector,  $k = 1, 2 \dots K$ ,  $L$  is the number of nodes in the hidden layer, the node serial number is  $l$ ,  $W_{jl}$  represents the connection weight between the  $j$ th neuron in the input layer and the  $l$  neuron in the hidden layer,  $T_{lk}$  represents the connection weight between the  $l$ th neuron in the hidden layer and the  $k$ th output neuron in the output layer, and the input of the  $l$ th node in the hidden layer is  $I_l$ . The output is  $y_l$ . The activation function is  $f(\cdot)$ . Then, the input and output of the  $l$ th neuron are expressed as

$$I_l = \sum_{j=1}^J W_{jl} x_j, \quad (11)$$

$$y_l = f(I_l).$$

Given  $N$  samples  $(\dot{X}_n, \dot{d}_n)$  ( $n = 1, 2, \dots, N$ ), the expected output vector of the input vector  $\dot{X}_n$  is  $\dot{d}_n$ . The objective function is defined as the sum of error squares between the expected output and the actual output during backpropagation.

$$E_n = \frac{1}{2} \sum_{k=1}^K [d_n(k) - O_n(k)]^2. \quad (12)$$

The total error of  $N$  samples is defined as

$$E = \frac{1}{2N} \sum_{n=1}^N \sum_{k=1}^K [d_n(k) - O_n(k)]^2. \quad (13)$$

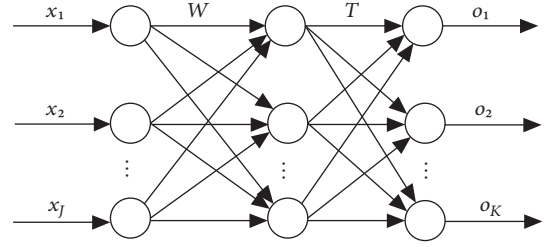


FIGURE 5: Topological structure of the BP neural network.

By adjusting the connection weight and threshold, the total error  $E$  is minimized and the weight changes along the negative gradient direction of the error function.

$$W_{jl}(t+1) = W_{jl}(t) - \eta \frac{\partial E}{\partial W_{jl}}, \quad (14)$$

where  $t$  is the number of iterations and  $\eta$  is the step size.

## 5. Fault Diagnosis Algorithm and Process

Based on the algorithm of the first two sections, a new method of short-circuit fault diagnosis for microgrid is proposed in this paper. The fault feature is extracted by wavelet packet decomposition, and the energy entropy is calculated. The multisource joint eigenvector is composed of voltage and current characteristic entropy, which is used as the BP network input and realizes fault diagnosis. Firstly, the time-frequency analysis of the three-phase current and voltage signals is carried out by wavelet packet decomposition. Then, the Shannon energy entropy is calculated as the signal feature vector, and the multisource joint eigenvector is composed of cross fusion and used as the input of the BP neural network for training and learning. The fault type and phase of microgrid can be accurately identified and diagnosed. The algorithm flow is shown in Figure 6.

500 groups of line fault sample data are randomly selected. Firstly, wavelet packet decomposition is used to analyze the signal time frequency, and then, Shannon energy entropy is used to extract the signal feature vector as the input signal of the BP neural network. The sample size for neural network training and learning is 80%. Another group of 10% data is used for verification, and the third group of 10% data is used for testing.

Using the BP neural network algorithm model, the multilayer feedforward network is trained according to the error backpropagation algorithm of microgrid fault output data, and then, through a large amount of sample learning, the input-output mode mapping relationship of the fault is stored, which can realize real-time and online mapping of fault information, making the complex nonlinear relationship in the output data samples become obvious. The accuracy of fault diagnosis is greatly improved, and the data error rate is reduced.



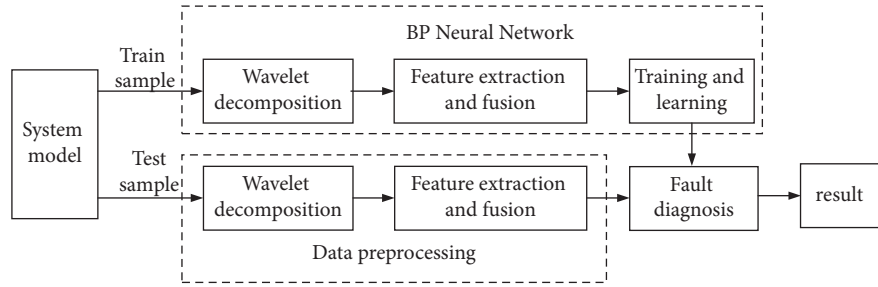


FIGURE 6: The flow chart of microgrid fault diagnosis.

## 6. Simulation Analysis and Experimental Verification

**6.1. Wavelet Packet Decomposition of Fault Current in Microgrid.** According to the experimental analysis and comparison, the frequency band distribution of 2-level wavelet packet decomposition is too wide and the resolution is low, while those of the 4-level wavelet packet decomposition and 3-level wavelet packet decomposition are the same, but the amount of calculation is significantly increased. Therefore, this paper chooses 3-level wavelet packet decomposition for the signal.

As an example, the step of extracting the wavelet packet energy entropy of the A-phase current signal is illustrated by the A-phase single-phase grounding short-circuit current of line WL3 away from the microgrid side. The waveform diagram of the single-phase grounding short-circuit current is shown in Figure 7.

Comparing Figures 7 and 5, it can be found that when the system is short-circuited, the current changes suddenly and has transient fault information. The db6 wavelet base is selected to further extract the effective fault information. The A-phase current signal is decomposed by 3-level wavelet multiresolution using formula (2), which can obtain the wavelet packet decomposition coefficient and wavelet reconstruction signal. The current decomposition signal is shown in Figure 8.

It can be seen from the detailed diagram that the first impulse current wave is received at 800 sampling points and the second impulse current wave is received at 3500 sampling points. Considering the period, it can be used as the basis for the fault location of microgrid.

The wavelet packet reconstruction of the short-circuit current and voltage wavelet signal shows that the fault signal contains rich nonstationary fault signal components, so the wavelet packet reconstruction signal will immediately have obvious fluctuations at the time of fault, which can be used as an important criterion to judge whether the fault occurs in the internal lines of the microgrid and to calculate the wavelet energy entropy value. The fault types and fault phases of the internal lines of the microgrid can be well prepared using the BP neural network.

**6.2. Extraction and Construction of Multisource Joint Eigenvector.** The energy entropy of 8 wavelet reconstruction signals of A-phase short-circuit current is calculated using

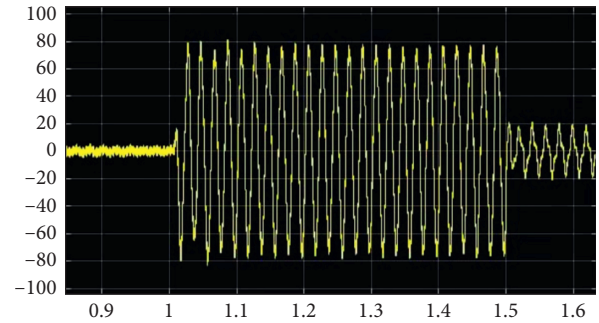


FIGURE 7: Short-circuit current of a single-phase earth fault.

Shannon entropy formula (8), and then, an eigenvector  $E$  is formed as follows:

$$E = [E_{30}, E_{31}, E_{32}, E_{33}, E_{34}, E_{35}, E_{36}, E_{37}]^T, \quad (15)$$

where  $E_{30}, E_{31}, E_{32}, E_{33}, E_{34}, E_{35}, E_{36}, E_{37}$  is the entropy of wavelet reconstruction signals. Because the wavelet packet Shannon entropy can detect small abnormal changes in the signal, when the signal-to-noise ratio is low, it can extract the effective weak signal and eliminate the noise very well. The smaller the entropy value is, the more orderly the signal is, and vice versa. It has little influence on the accuracy of judgment with the amount of calculation being multiplied in the process of fault diagnosis and recognition using BP neural network; at the same time, the high-frequency signal component is more than one order of magnitude smaller than the low-frequency signal component, so  $E_{30}, E_{31}, E_{32}, E_{33}$  can be taken as the input of BP neural network.

$$E = [E_{30}, E_{31}, E_{32}, E_{33}]^T. \quad (16)$$

In the same way, the other two-phase current and voltage signals are processed and 16-dimensional wavelet energy entropies are obtained. The multisource fault information is fused by the interval cross mode to form the feature vector for fault diagnosis. The fusion mode is shown in Figure 9.

Some multisource feature eigenvectors are shown in Table 1.

Compared with the normal state energy entropy, some results are shown in Figure 10.

The energy entropy  $E_{31}, E_{32}, E_{33}$  of phase A is surely increased when A phase shows ground fault but B and C phases almost remain unchanged; the energy entropy  $E_{31}, E_{32}, E_{33}$  of two phases A and B are significantly increased when two-phase short circuit or two-phase grounding short

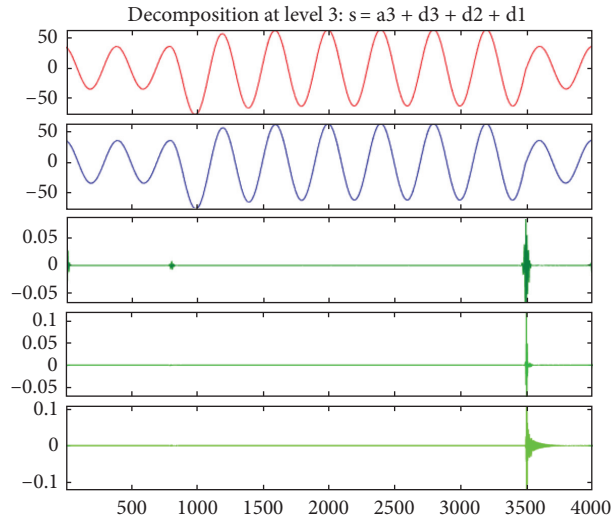


FIGURE 8: Wavelet packet decomposition of a single-phase grounding short-circuit current.

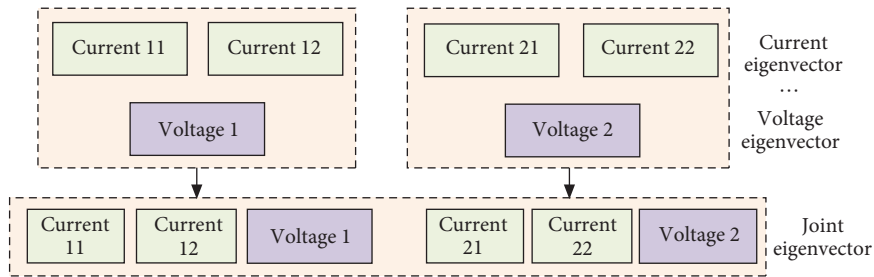


FIGURE 9: Association diagram of the joint fault feature vector.

TABLE 1: Some multisource feature eigenvectors.

Fault type	Fault position (%)	Feature eigenvector
AG	10	16.110, 4.3607, 0.0810, 0.1037, 23.616, 0.7045, 0.1075, 0.0562, 21.344, 0.5773, 0.0865, 0.0411, 56.320, 0.8069, 0.5211, 0.04385
AG	50	84.916, 0.4371, 0.0114, 0.0135, 22.398, 0.4777, 0.1000, 0.0537, 21.780, 0.3528, 0.0795, 0.0390, 125.102, 1.0972, 0.3917, 0.0846
ABG	10	17.934, 3.8371, 0.1305, 0.4891, 16.407, 3.3080, 0.2022, 0.1636, 15.357, 0.6490, 0.0887, 0.0665, 51.105, 0.9281, 0.5371, 0.0721
ABG	50	11.619, 1.2852, 0.0338, 0.1220, 10.578, 1.2189, 0.1338, 0.1893, 19.462, 0.4162, 0.0817, 0.0453, 59.219, 1.036, 0.9024, 0.06901
ABG	85	86.968, 0.2493, 0.0066, 0.0189, 80.033, 0.7853, 0.1145, 0.1464, 21.326, 0.3413, 0.0782, 0.0403, 129.264, 1.5973, 0.7941, 0.1026
AB	10	14.133, 4.2099, 0.0692, 0.1754, 14.353, 4.6465, 0.1665, 0.2262, 25.090, 0.3239, 0.0775, 0.0377, 22.683, 0.9251, 0.0529, 0.0904
AB	85	76.814, 1.0975, 0.0360, 0.1069, 76.832, 1.5341, 0.1328, 0.1570, 25.102, 0.3247, 0.0777, 0.0378, 110.469, 2.1963, 0, 9625, 0.0805
ABC	10	18.660, 1.1549, 0.0750, 0.4304, 17.224, 2.7162, 0.1434, 0.2369, 17.290, 2.7340, 0.1746, 0.5255, 59.132, 1.6938, 0.9934, 0.0982
ABC	70	10.779, 0.3454, 0.0274, 0.0971, 10.100, 1.7274, 0.1197, 0.1159, 10.127, 1.0555, 0.1089, 0.1373, 49.669, 1.0394, 0.6392, 0.0872
ABCG	10	18.615, 3.6540, 0.0415, 0.3515, 17.163, 3.3278, 0.1675, 0.3161, 17.579, 4.5462, 0.1377, 0.5702, 38.672, 0.9821, 0.0923, 0.0639
ABCG	70	10.766, 1.2417, 0.0176, 0.0644, 10.077, 1.9576, 0.1250, 0.1326, 10.210, 1.1433, 0.0915, 0.0571, 44.971, 0.7953, 0.0926, 0.0481
Normal		25.098, 0.0126, 0.0012, 0.0015, 25.088, 0.4485, 0.0973, 0.0522, 25.090, 0.3239, 0.0775, 0.0377, 64.925, 1.0369, 0.9835, 0.0485



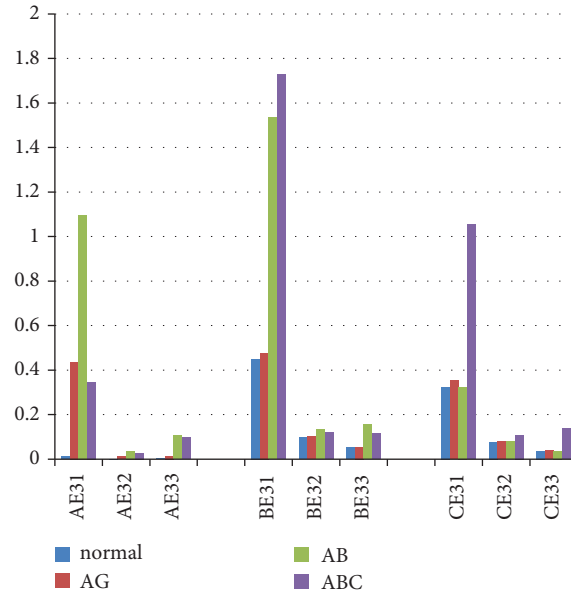


FIGURE 10: Energy entropy comparison.

circuit occurs; the energy entropy  $E_{31}$ ,  $E_{32}$ ,  $E_{33}$  of three phases are all improved when A, B, and C three-phase short-circuit occurs. This can be used as a basis for fault type identification and fault phase judgment in microgrid.

**6.3. Training Results of BP Neural Network.** Three-phase current and voltage eigenvalues constitute a multisource fault eigenvector which is taken as the input vector of the BP neural network. The fault signal of microgrid is decomposed and reconstructed with a wavelet, and the number of input neuron nodes is set as 16. The state of the three phases and the neutral line are taken as output vectors, so the number of output neuron nodes of the neural network is set as 4. The output value of 1 indicates that the corresponding line is at fault or the fault phase is grounded, and 0 indicates that the corresponding line has no fault. The number of hidden layer neurons affects the training results of the model. The training accuracy is poor if too few nodes are selected; the training time and step size are relatively large when there are too many nodes, and it is easy to fit. This is verified by experiments, and combined with the empirical formula  $L = \sqrt{J + K} + \alpha$ , this paper chooses 10 hidden layer nodes. The BP network stops the model training and learning when the training error meets the given requirements. The training curve is shown in Figure 11.

The fitting degree curve of neural network training is shown in Figure 12. Through observation, the fitting degree of neural network is high, the fitting degree of training and testing is more than 0.9, and the fitting degree of verification is also 0.88. Thus, BP neural network can accurately diagnose and identify faults.

**6.4. Hardware in the Loop Simulation Test.** Because the actual microgrid systems are built outdoors, where the natural conditions are bad and the system structure is complex,

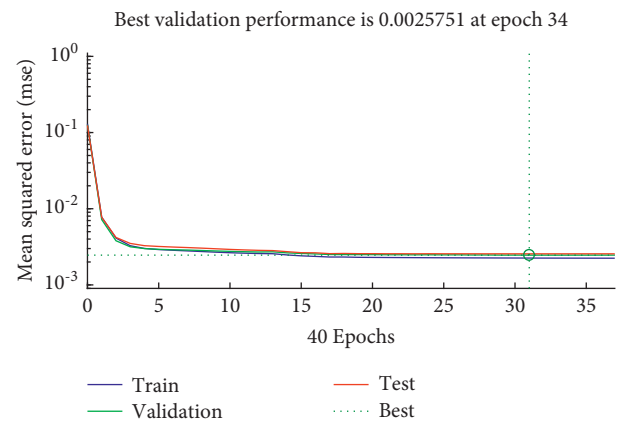


FIGURE 11: Training curve.

including a large number of power electronic device load units, the voltage level is high and dangerous. However, the construction of a traditional electrical laboratory has high construction costs and site requirements. It is difficult to simulate and research some special conditions such as faults considering safety and equipment maintenance. Therefore, the hardware in the loop simulation technology is applied to the experimental environment based on the existing equipment and conditions of the laboratory. The hardware in the loop simulation platform is composed of PXI hardware and StarSim HIL and StarSim RCP software developed by Yuanquan energy to realize microgrid fault identification and diagnosis. The hardware in the loop platform is shown in Figure 13, the system framework is shown in Figure 14, and the microgrid interface is shown in Figure 15.

**6.5. Fault Diagnosis and Identification.** The multisource feature vectors of microgrid test samples are brought into the trained BP neural network to diagnose and identify the fault

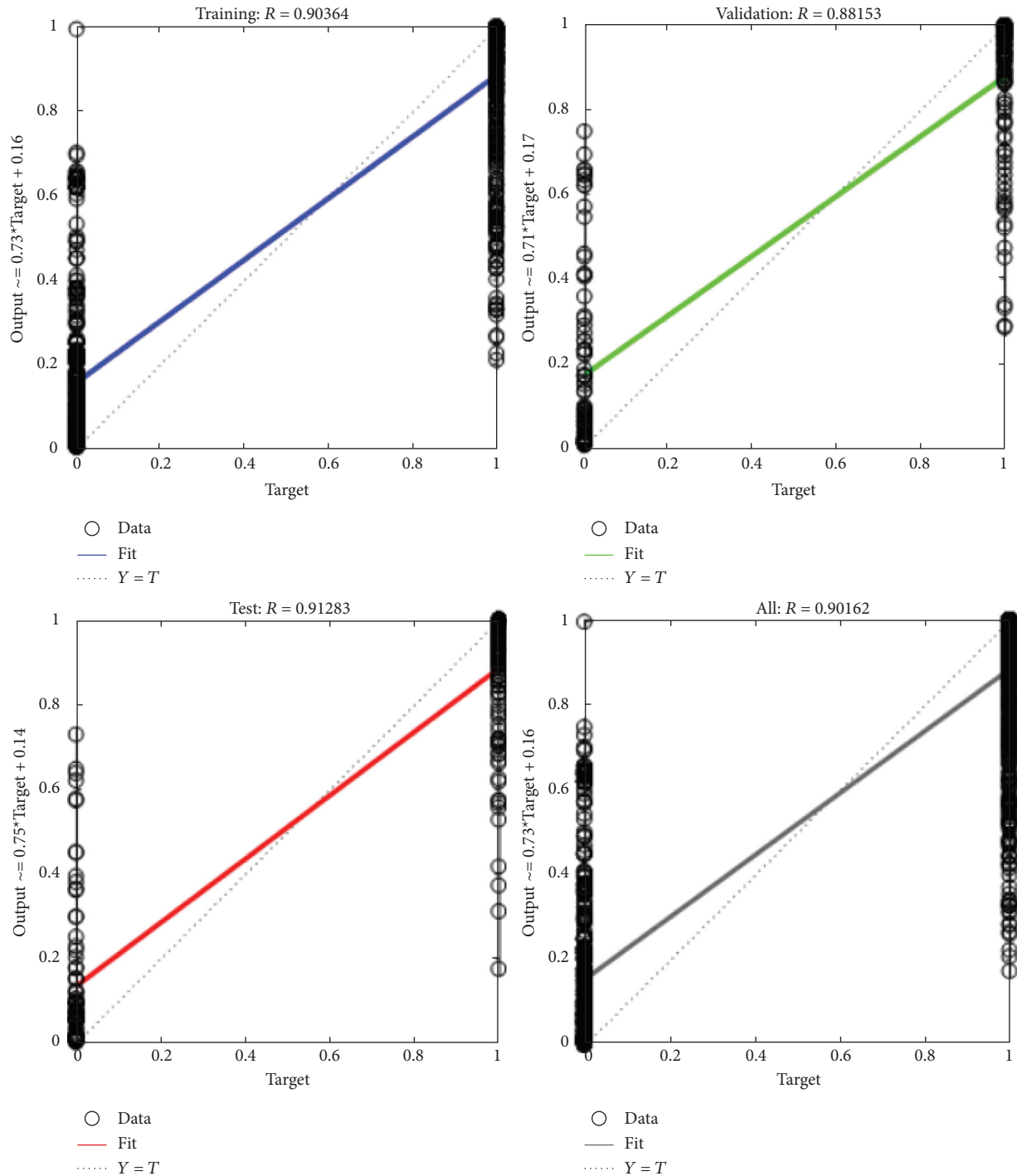


FIGURE 12: Fitting curve.

type and phase. The actual output signal of the test sample is compared with the expected output signal. Some test samples and their fault diagnosis results are shown in Table 2.

The BP neural network method based on big data analysis extracts and abstracts the features of fault data and strengthens the spatiotemporal correlation of heterogeneous data to realize the prediction of fault type and phase. At the same time, the BP neural network method based on big data obtains accurate data information by constructing many

invisible models and performing a large amount of data analysis and training. It can be seen from Table 2 that the trained BP neural network model can accurately and effectively identify the fault type and fault phase and the error between the actual output and the expected output value meets the requirements of fault diagnosis. The test results show that two samples have fault diagnosis among 50 samples; the accuracy rate reaches 96%, which meets the requirements for intelligent fault diagnosis of microgrid lines.



FIGURE 13: The hardware of the loop platform.

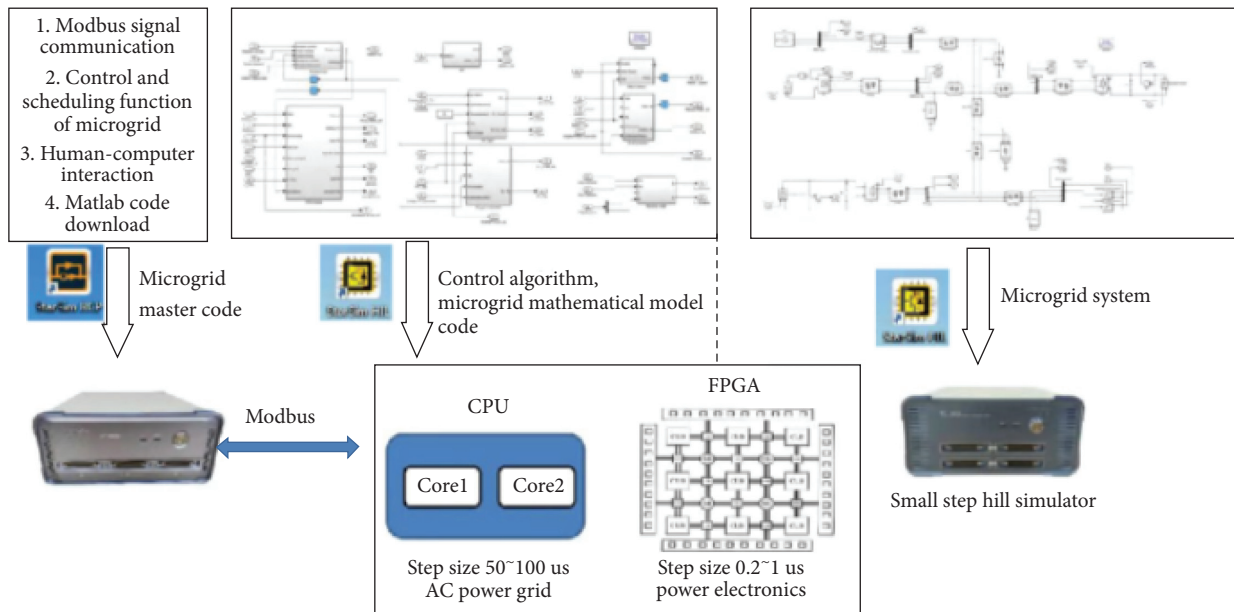


FIGURE 14: The system framework.

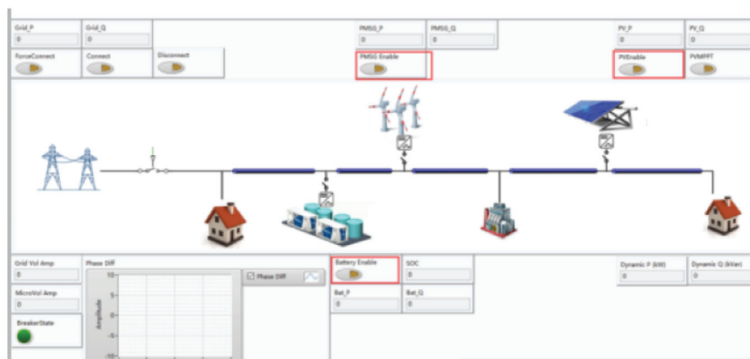


FIGURE 15: The microgrid interface.

TABLE 2: Part of the test samples and fault diagnosis results.

Fault type	Expected output	Test sample	Actual output
AG	1, 0, 0, 1	60.968, 0.062, 0.007, 0.034, 22.730, 0.451, 0.098, 0.054, 22.653, 0.327, 0.078, 0.039, 89.361, 1.804, 0.9735, 0.0639	0.973, 0.017, 0.015, 0.987
AC	1, 0, 1, 0	14.410, 1.223, 0.076, 0.414, 25.088, 0.449, 0.097, 0.052, 14.183, 1.532, 0.151, 0.450, 49.035, 1.329, 0.405, 0.0579	0.998, 0.027, 1.045, 0.028
ACG	1, 0, 1, 1	91.732, 1.770, 0.041, 0.104, 20.424, 0.516, 0.099, 0.056, 96.708, 1.272, 0.089, 0.071, 120.488, 1.409, 0.1088, 0.053	1.009, 0.025, 0.998, 1.012

## 7. Conclusions

Aiming at the characteristics of real-time transmission, fast updating, and large-scale fault information data of microgrid, a microgrid fault diagnosis and analysis technology supported by big data is proposed in this paper. This technology combines Rayleigh entropy, wavelet packet decomposition method, and BP neural network to extract the fault feature vector of microgrid. The BP neural network method based on big data strengthens the spatiotemporal correlation of heterogeneous data to realize the prediction of fault type and phase. The BP neural network method based on big data obtains accurate results by constructing many invisible models and a large amount of data training. The experimental results show that the fault diagnosis and analysis technology based on big data support proposed in this paper has an accuracy of 96%, which fully meets the needs of engineering practical application. However, due to the complex topology and many fault types of microgrid, only five line fault types are considered in this paper. Therefore, in the next step, the technical method proposed in this paper needs to be applied to other fault diagnosis to verify the universality of the technical method proposed in this paper.

## Data Availability

The datasets used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was partially supported by Zhejiang Province Visiting Scholar Program: Theory and Practice of Event Driven Control in Microgrid (FX2021069).

## References

- [1] K. P. Rajendra and S. V. Kumar, "Multi Agent System Based Intelligent Fault Diagnosis with Fault Current Limiter for Microgrid," in *Proceedings of the National Power Systems Conference*, pp. 1–6, IEEE, Bhubaneswar, India, December 2016.
- [2] Q. Lu, Y. Z. Ye, and C. D. Jiang, "Fault diagnosis of microgrid system based on wavelet singular entropy and SOM neural network," *Journal of Shandong University*, vol. 47, no. 5, pp. 118–129, 2017.
- [3] T. J. Sun, L. P. Qu, and H. S. Guan, "Power grid fault diagnosis based on intelligent optimization fuzzy Petri net," *Control Engineering*, vol. 28, no. 1, pp. 127–134, 2021.
- [4] Q. Kang, Y. Luo, X. J. Lu, and P. Gao, "Simulation of microgrid fault characteristics based on converter control strategy," *Power system protection and control*, vol. 47, no. 2, pp. 147–153, 2019.
- [5] P. Hu, "Analysis and Research on short circuit fault of distribution line in microgrid system," *Machinery Manufacturing*, vol. 635, no. 55, pp. 111–113, 2017.
- [6] C. D. Jiang, L. Qiu, and Y. Z. Ye, "Short circuit fault diagnosis of microgrid based on Shannon entropy and BP neural network," *Southern Agricultural Machinery*, vol. 24, no. 2, pp. 14–16, 2020.
- [7] Y. J. Yang, Y. Chen, and K. Yang, "Fault diagnosis of microgrid line based on wavelet packet analysis and neural network," *Automatica e Instrumentacion*, vol. 38, no. 9, pp. 65–74, 2017.
- [8] Y. J. Yang and F. Mao, "Fault diagnosis of optical storage microgrid line based on RTDS and neural network," *Renewable Energy*, vol. 36, no. 7, pp. 1010–1016, 2018.
- [9] S. Liu and D. Yang, "Identification and detection algorithm of electric energy disturbance in microgrid based on wavelet analysis and neural network," *EURASIP Journal on Wireless Communications and Networking*, vol. 1, pp. 1–9, 2021.
- [10] S. Jamali, S. Ranjbar, and A. Bahmanyar, "Identification of faulted line section in microgrids using data mining method based on feature discretisation," *International Transactions on Electrical Energy Systems*, vol. 30, no. 6, pp. 2104–2119, 2020.
- [11] C. D. Jiang, W. Yang, and Y. Guo, "Nonlocal means two dimensional histogram-based image segmentation via minimizing relative entropy," *Entropy*, vol. 20, no. 10, pp. 826–837, 2018.
- [12] J. H. Chen, Y. J. Gao, and Y. W. Jin, "Microgrid fault diagnosis method based on MODWT and BP neural network," *Journal of Shanghai University of Electric Power*, vol. 37, no. 1, pp. 57–60, 2021.

## Research Article

# Production Task Allocation Decision Based on Cloud Robot Cell-Line

Yi Du <sup>1</sup>, Yongqiang Wang,<sup>1</sup> Jialin Wang,<sup>2</sup> and Tianpeng Zhang<sup>3</sup>

<sup>1</sup>Anyang Institute of Technology School of Electronic Information and Electrical Engineering, Anyang 455000, China

<sup>2</sup>Anyang Branch, China Mobile Tietong, Anyang 455000, China

<sup>3</sup>Hangzhou Normal University, Hangzhou 310000, China

Correspondence should be addressed to Yi Du; 20170020@ayit.edu.cn

Received 11 January 2022; Revised 14 February 2022; Accepted 16 February 2022; Published 24 March 2022

Academic Editor: Daqing Gong

Copyright © 2022 Yi Du et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Cloud manufacturing is a new service-oriented efficient and low-consumption agile manufacturing mode integrating information, manufacturing, Internet of Things, and other technologies. One of the key decisions of production enterprises is production task allocation based on cloud robot cell-line, which determines the efficiency and flexibility of the production system and affects various production links, such as job-shop logistics, production planning, and production scheduling. This paper explores the production task allocation, from the angle of the optimal combination of cloud manufacturing resources. First, a mathematical model was established, based on the transport cost of different sub-tasks and the tardiness cost of product delivery, and solved by quantum firefly algorithm (QFA). Next, QFA was proved superior to traditional firefly algorithm (FA), improved FA, and the FA optimized by cat swarm optimization (CSO-FA), in terms of time complexity and spatial complexity. The research enriches the theory and methodology of allocating operation-level cloud manufacturing resources based on cloud robot cell-line and provides decision support to manufacturers, which want to implement operation-level allocation of cloud manufacturing resources based on cloud robot cell-line.

## 1. Introduction

Big manufacturers like China are improving the intelligence level of the manufacturing industry. Cloud intelligence can increase the utilization efficiency of production lines, lower the production cost, and avoid repeated investment and construction, contributing to the improvement of our living standards.

Unlike traditional flow lines, the cell-line combines a few similar stations into a novel type of production line. There are two advantages of the cell-line: First, it is of the same production efficiency as a conveyor belt flow line. Second, the flexible production of production units adapts well to the market needs of various small batches of products. Therefore, the cell-line is regarded as an ideal production mode to realize massive customization. Besides, the cell-line provides a green and economic production method.

Scholars at home and abroad have explored robotic automation production lines extensively. Focusing on the variable process pathway, Tian et al. [1] established a minimum delay time function and constructed a scheduling module, in the light of collaborative, differentiated agents. In addition, agents were adopted to realize global optimization, aiming to minimize the makespan. Zacharia and Nearchou [2] summarized the features of job-shop scheduling with multiple robotic manufacturing units, created a mathematical model for minimizing the makespan of the problem, and solved the model effectively with a hybrid genetic algorithm. Liu et al. [3] utilized the minimum completion time to solve the flexible job-shop scheduling with loading/unloading robots, established a mathematical model of the problem, and searched for the scheduling results with a tabu search and greedy algorithm. Miyake [4] used the minimum production cycle to solve the multi-process job-shop scheduling with transport time, constructed a nonlinear

mathematical model, and solved the model with improved genetic algorithm, realizing batch scheduling. Du et al. [5] explored the reconstruction and scheduling of multi-stage variable-batch production lines in labor-intensive enterprises, adopted different algorithms for production lines of different scales, and compared the applicable scopes of two joint optimization algorithms with examples. Wu et al. [6] studied the optimal control of event-driven conveyor belt feeding and processing system.

There are some limitations that present opportunities for future research. Our Stackelberg game modeled the operator as the leader and the suppliers as the followers. It might be beneficial for practical applications to explore other power structures. In future research, a supplier-led power structure (e.g., a large-scale enterprise building a platform) could be an important extension of this model. Moreover, we studied two types of substitutable cloud services that can satisfy client demands. In reality, there might be multiple types of cloud services that meet the same requirement. Future research should consider the case of multiple types of substitutable cloud services in the CMfg system.

Cao et al. [7] stated that there are some limitations that present opportunities for future research. Our Stackelberg game modeled the operator as the leader and the suppliers as the followers. It might be beneficial for practical applications to explore other power structures. In future research, a supplier-led power structure (e.g., a large-scale enterprise building a platform) could be an important extension of this model. Moreover, we studied two types of substitutable cloud services that can satisfy client demands. In reality, there might be multiple types of cloud services that meet the same requirement. Future research should consider the case of multiple types of substitutable cloud services in the CMfg system. Carlucci et al. [8] observed that a key limitation of the research is the use of homogeneous resources shared in the network. Future research could take into account different categories of resources that are characterized by different rules in order to determine the win or lose state. Another important research issue to further investigate is the dynamic composition of the network based on the obtained performance, i.e., how the participant in the network can decide to enter or leave the network.

To sum up, there are very few studies on cell-line cloud manufacturing resource scheduling based on cloud robots, which could avoid resource waste, repetitive building of production lines, and low utilization of units. This paper explores production task allocation from the angle of the optimal combination of cloud manufacturing resources. Specifically, a mathematical model was established based on the transport cost of different sub-tasks and the tardiness cost of product delivery, and solved by quantum firefly algorithm (QFA). The effectiveness of the model and algorithm was demonstrated through example analysis.

At present, there are many researches on the single technology of cloud manufacturing. With the in-depth research and application of cloud manufacturing, the cloud manufacturing service platform oriented to the whole life cycle of products will become the next research hotspot. New product development is an important life cycle stage of

complex electronic products. There are very few researches on cloud manufacturing resource scheduling of the Japanese unit production line based on cloud robot. However, the results of researches can avoid the waste of resources, repeated line construction, and low utilization rate of units.

The driving effect of electronic information industry to China's economic growth is showing a sustained and scale of growth. Moreover, the promotion and penetration effect to the upgrading of traditional industrial structure gradually increases. The high-tech industry represented by the electronic information industry has developed far faster than the traditional industry, and the export of high-tech products has shown a rapid growth trend. By observing the development of electronic information industry and the track of technological upgrading, it is concluded that China has a strong and obvious comparative advantage in processing and assembly, and the manufacturing of electronic information products in the world is forming a tendency to be concentrated in some areas of China. First, due to the epidemic, foreign orders have increased, and some manufacturers have blindly expanded production scale, resulting in serious excess production capacity. However, some manufacturers have insufficient orders and idle production resources. Second, the processing technology of small batch products is more and more complex and often the production resources of an enterprise cannot meet the requirements of product processing. The cloud platform can contact enterprises in multiple regions to complete manufacturing resource scheduling tasks together.

## 2. Problem Formulation

In 2010, James first put forward the definition of "cloud" robot; it points out that the cloud robotics is robot to store and retrieve information, such as a robot with a camera around the photo upload cloud server, a search on the server, find the similar photos, to plan the robot path tracing, and store the pictures, as search information for the back of the machine. All robots can share this library to reduce the developer effort and time.

Suppose there are requests for  $I$  types of products on the cloud platform. Each product request contains  $N$  sub-tasks  $F_i$ ,  $i = \{1, 2, \dots, N\}$  to be processed on  $k$  intelligent cloud scheduling production lines. Each sub-task contains  $P_i$  operations, and  $K_i$  optional production lines. The production task can be expressed as follows:

$$F = \min(f_1, f_2), \quad (1)$$

where  $f_1$  is the minimum processing time of the manufacturing resources of a task, which depends on the sub-task with the longest processing time:

$$f_1 = \min \left( \max_{t=1}^T \sum_{i=1}^I \sum_{k=1}^K (e_{tik} \varepsilon_{ti} L_{tik} + x_{k,k+i} q_{tik} h_{ti}) \right), \quad (2)$$

where  $e_{ikt}$  is a 0–1 variable (if  $e_{ikt} = 1$ , product  $i$  is processed through  $t$  stages on production line  $k$ ; if  $e_{ikt} = 0$ , product  $i$  is not processed through  $t$  stages on production line  $k$ );  $\varepsilon_{ti}$  is



the number of product  $i$  through  $t$  stages;  $l_{ikt}$  is the processing time of product  $i$  through  $t$  stages on production line  $k$ ;  $q_{tik}$  is the transport time of 1,000 products  $i$  from production line  $k$  to production line  $k+1$ ;  $x_{k,k+i}$  is a 0-1 variable (if  $x_{k,k+i} = 1$ , product  $i$  is transported from production line  $k$  to production line  $k+1$  through  $t$  stages; if  $x_{k,k+i} = 0$ , product  $i$  is not transported from production line  $k$  to production line  $k+1$  through  $t$  stages);  $h_{ti}$  is the multiple of 1,000 products  $i$  through  $t$  stages.

$f_2$  is the minimum cost of all sub-tasks of manufacturing resource processing:

$$f_2 = \min \left( \max_{t=1}^T \sum_{i=1}^I \sum_{k=1}^K (e_{tik} \varepsilon_{ti} C_{tik} + (f_{itk} - s_{itk}) * d_{it}) \right), \quad (3)$$

where  $C_{tik}$  is the manufacturing cost of product  $i$  through  $t$  stages on production line  $k$ ;  $f_{itk}$  is the processing time of product  $i$  through  $t$  stages on production line  $k$  (including transport time);  $s_{itk}$  is the delivery period of product  $i$  through  $t$  stages on production line  $k$ ; and  $d_{itk}$  is the tardiness cost constraint of product  $i$  through  $t$  stages.

The processing of the current operation of a sub-task should not begin before the completion of the previous operation:

$$\sum_{k=1}^{M_{ij}} e_{ijk} = 1, \quad (4)$$

where  $t_{\text{end},j}$  and  $t_{\text{start},i(i+1)}$  are the completion time of the previous operation and the start time of the current operation, respectively.

Operation is the minimum granularity of scheduling. Each operation can only occupy one logic manufacturing resource.

### 3. Algorithm Design

The QFA is designed in the following steps [9-14]. First, a swarm of  $h$  quantum fireflies is generated. Each quantum firefly has a quantum position and a position. The position of each quantum firefly is represented by a  $D$ -dimensional numerical string  $\{0, 1\}$ , with  $D$  being the dimensionality of the solution space. In the  $t$ -th iteration, the quantum position of the  $i$ -th quantum firefly can be expressed as follows:

$$v_i^t = \begin{bmatrix} \alpha_{i1}^t & \alpha_{i2}^t & \dots & \alpha_{iD}^t \\ \beta_{i1}^t & \beta_{i2}^t & \dots & \beta_{iD}^t \end{bmatrix}, \quad (5)$$

where  $(\alpha_{ij}^t)^2 + (\beta_{ij}^t)^2 = 1$  ( $i = 1, 2, \dots, D$ ). The quantum positions  $\alpha_{ij}^t$  and  $\beta_{ij}^t$  satisfy  $0 \leq \alpha_{ij}^t \leq 1$ , and  $0 \leq \beta_{ij}^t \leq 1$ . At the beginning, all the quantum positions of the quantum fireflies are initialized as 1. For simplicity, the quantum position of the  $i$ -th quantum firefly in the  $t$ -th iteration can be denoted as  $v_i^t = (v_{i1}^t, v_{i2}^t, \dots, v_{iD}^t)$  ( $i = 1, 2, \dots, h$ ). The position of a quantum firefly can be derived from the measured quantum position. The position of the  $i$ -th quantum firefly can be expressed as  $x_i^t = (x_{i1}^t, x_{i2}^t, \dots, x_{iD}^t)$  ( $i = 1, 2, \dots, h$ ). The global optimal position of the  $i$ -th

quantum firefly in the  $t$ -th iteration can be expressed as  $p_g^t = (p_{g1}^t, p_{g2}^t, \dots, p_{gD}^t)$ .

According to the fluorescein update rule, the fitness  $F$  ( $p_i^t$ ) of quantum firefly  $i$  ( $i = 1, 2, \dots, h$ ) at the local optimal position in the  $t$ -th iteration can be converted into fluorescein  $L_i(t)$ :

$$L_i(t) = (1 - \gamma)L_i(t-1) + \varepsilon F(p_i^t), \quad (6)$$

where  $\gamma \in [0, 1]$  is the disappearance rate of fluorescein, which gradually weakens with the growing distance and media absorption;  $\varepsilon$  is the update rate of fluorescein. The  $i$ -th quantum firefly obtains the learning neighborhood by certain rules. The quantum fireflies are selected from the neighborhood based on the fluorescein value and position similarity. In the  $t$ -th iteration, the  $i$ -th quantum firefly obtains the learning neighborhood by the following equation:

$$N_i(t) = \{q | D - d_{iq}(t) \leq r_i(t) \text{ and } L_i(t) \leq L_q(t)\}, \quad (7)$$

where  $N_i(t)$  is the set of labels of the learning neighborhood of the  $i$ -th quantum firefly;  $d_{iq}(t)$  is the distance between the local optimal positions of the  $i$ -th and  $q$ -th quantum fireflies;  $r_i(t)$  is the radius of the dynamic decision domain of the  $i$ -th quantum firefly; and  $L_q(t)$  is the fluorescein value of the  $q$ -th quantum firefly in the  $t$ -th iteration. The number of quantum fireflies in the learning neighborhood equals the number of labels in the label set of that neighborhood in the current iteration.

In each iteration, each quantum firefly selects its moving direction according to the quantum fireflies in the neighborhood. In the  $t$ -th iteration, the probability for the  $i$ -th quantum firefly to move to the  $q$ -th quantum firefly in its neighborhood can be expressed as

$$P_{iq}' = \frac{L_q(t) - L_i(t)}{\sum_{j=N_i(t)} [L_j(t) - L_i(t)]}. \quad (8)$$

In each iteration, if the learning neighborhood of the  $i$ -th quantum firefly is empty, then the evolution of the 1-th dimensional quantum position of that quantum firefly can be expressed as follows:

$$v_{il}^{t+1} = \begin{cases} N v_{il}^t \theta_{il}^{t+1} = 0 \text{ and } \mu_{il}^{t+1} < c_1, \\ \text{abs}[U(\theta_{il}^{t+1}) v_{il}^t] \end{cases}, \quad (9)$$

where quantum rotation angle  $\theta_{il}^{t+1} = e_1 (p_{il}^t - x_{il}^t) + e_2 (p_{gl}^t - x_{il}^t)$  ( $i = 1, 2, \dots, h; l = 1, 2, \dots, D$ ;  $e_1$  and  $e_2$  are the degree of influence of local optimal position and global optimal position on quantum rotation angle, respectively;  $\mu_{il}^{t+1}$  is a random number uniformly distributed in  $[0, 1]$ ;  $c_1$ , a constant in  $[0, 1/D]$ , is the mutation probability of the quantum position at the quantum rotation angle of 0;  $\text{abs}(\cdot)$  is the absolute value that limits each dimension of the quantum position in  $[0, 1]$ ;  $U(\theta_{il}^{t+1}) = \begin{bmatrix} \cos \theta_{il}^{t+1} & -\sin \theta_{il}^{t+1} \\ \sin \theta_{il}^{t+1} & \cos \theta_{il}^{t+1} \end{bmatrix}$  is the quantum rotation gate; and  $\bar{N} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$  is the quantum NOT gate.

If the learning neighborhood of the  $i$ -th quantum firefly is not empty, each quantum firefly labeled  $z$  in the neighborhood satisfies  $z = \arg \max_{q \in N_{il}} \{ \overline{P}_{iq}(t) \}$ . In the  $t$ -th iteration, then the evolution of the  $l$ -th dimensional quantum position of the  $i$ -th quantum firefly can be expressed as follows:

$$v_{il}^{t+1} = \begin{cases} \overline{N}v_{il}^t & \theta_{il}^{t+1} = 0 \text{ and } d\mu_{il}^{t+1} < c_2 \\ \text{abs}[U(\theta_{il}^{t+1})v_{il}^t] & \end{cases}, \quad (10)$$

where quantum rotation angle  $\theta_{il}^{t+1} = e_3(p_{il}^t - x_{il}^t) + e_4(p_{zl}^t - x_{il}^t) + e_5(p_{gl}^t - x_{il}^t)$ ;  $i = 1, 2, \dots, h$ ;  $l = 1, 2, \dots, D$ ;  $p_{zl}^t$  is the  $l$ -th dimension of the local optimal position, i.e., the position with the highest fluorescein, in the learning neighborhood of the  $i$ -th quantum firefly;  $e_3, e_4$ , and  $e_5$  are the degree of influence of the local optimal position of the  $i$ -th quantum firefly, the local optimal position, i.e., the position with the highest fluorescein, in the learning neighborhood of the  $i$ -th quantum firefly, and the global optimal position of the  $i$ -th quantum firefly on quantum rotation angle, respectively;  $c_2$ , a constant in  $[0, 1/D]$ , is the mutation probability of the quantum position at the quantum rotation angle of 0. The position of the  $i$ -th quantum firefly can be derived from the measured quantum position:

$$x_{il}^{t+1} = \begin{cases} 1, \eta_{il}^{t+1} > (\alpha_{il}^{t+1})^2 \\ 0, \eta_{il}^{t+1} \leq (\alpha_{il}^{t+1})^2 \end{cases}, \quad (11)$$

where  $l = 1, 2, \dots, D$ ;  $\eta_{il}^{t+1} \in [0, 1]$  is a random number obeying uniform distribution;  $(\alpha_{il}^{t+1})^2$  is the probability of zero appearing at quantum position  $v_{il}^{t+1}$ .

The radius of the dynamic decision domain of the  $i$ -th quantum firefly can be updated by the following expression:

$$r_i(t+1) = \min\{R_s, \max[0, r_i(t) + \zeta\{n_t - \text{size}[N_i(t)]\}]\}, \quad (12)$$

where  $\zeta$  is a constant update rate of the dynamic decision domain;  $R_s \geq r_i(t)$  is a constant perception domain;  $\min$  and  $\max$  are the minimum and maximum functions, respectively;  $n_t$  is the parameter controlling the number of quantum fireflies in the learning neighborhood; and  $\text{size}[N_i(t)]$  is the number of quantum fireflies in the learning neighborhood of the  $i$ -th quantum firefly.

For the cloud scheduling of production line, the fitness function of the QFA should minimize the production time of the production line. After computing the maximum production time and cost of cloud scheduling, the minimization of the two values is adopted as the optimization goal. Then, the fitness function of the current position  $x_i^t = (x_{i1}^t, x_{i2}^t, \dots, x_{iD}^t)$  ( $i = 1, 2, \dots, h$ ) of the  $i$ -th firefly can be expressed as follows:

$$F(x_i^t) = \begin{cases} -MSLL(x_i^t), & \text{If } crat \leq erat \\ -\rho \cdot MSLL(x_i^t) & \end{cases}, \quad (13)$$

where  $MSLL(x_i^t)$  is the maximum production time and cost of cloud scheduling constructed by  $x_i^t$ ;  $\rho \leq 1$ ;  $crat$  and  $erat$

are the calculated and expected coverage rates, respectively [15–19].

Based on the QFA, the production time and cost of the cloud manufacturing production line can be minimized through the following steps:

*Step 1.* Establish the mathematical model of the production time and cost of cloud computing production line, determine the key parameters of the QFA corresponding to the production time and cost, initialize the quantum positions of quantum fireflies, and measure their positions.

*Step 2.* Map the position of each quantum firefly to the corresponding parameters of the established model, import these parameters to the fitness function, compute the fitness at the position of the quantum firefly, and determine the local and global optimal positions according to the fitness.

*Step 3.* Update the fluorescein value and learning neighborhood of each quantum firefly, according to the fitness at the local optimal position of that quantum firefly.

*Step 4.* Change the quantum position of each quantum firefly with new data.

*Step 5.* Change the radius of the dynamic decision domain of each quantum firefly with new data.

*Step 6.* Compute the fitness of each firefly at the new position by the fitness function and determine the local and global optimal positions according to the fitness again.

*Step 7.* If the maximum number of iterations is reached, execute Step 8; otherwise, return to Step 3.

*Step 8.* Output the global optimal position [20–25]. (Flow chart as Figure 1)

The test example and calculation results are shown in Tables 1 and 2, respectively.

## 4. Example Analysis

Two swarms were initialized in our example. For convenience, each type of the product can be used after completion one production cycle. The initial swarms 1 and 2 were combinations of ones in each row of Table 1. Our algorithm IFA was compared with traditional firefly algorithm (FA), improved FA, and the FA optimized by cat swarm optimization (CSO-FA). The swarm size for all these intelligent optimization algorithms was fixed at 50 (Set in the program). The maximum number of iterations was set to 350 (Set in the program), the expected coverage rate to 0.7 (In (13),  $erat$  is the expected filling rate),  $\rho = 0.001$  (In Formula (13)  $\rho \leq 1$ ), the update rate of dynamic decision domain to 0.8 (Formula (12)  $\zeta$  is the update rate of the dynamic decision domain), the initial fluorescein value to 5,  $e_1 = 0.06$ ,  $e_2 = 0.03$ ,  $e_3 = 0.06$ ,  $e_4 = 0.03$ ,  $e_5 = 0.01$ ,  $\theta_{il}^{t+1} = e_3(p_{il}^t - x_{il}^t) + e_4(p_{zl}^t - x_{il}^t) + e_5(p_{gl}^t - x_{il}^t)$ ;  $i = 1, 2, \dots, h$ ;  $l = 1, 2, \dots, D$ ;  $p_{zl}^t$  and  $e_2$  are the degree of influence of local optimal position

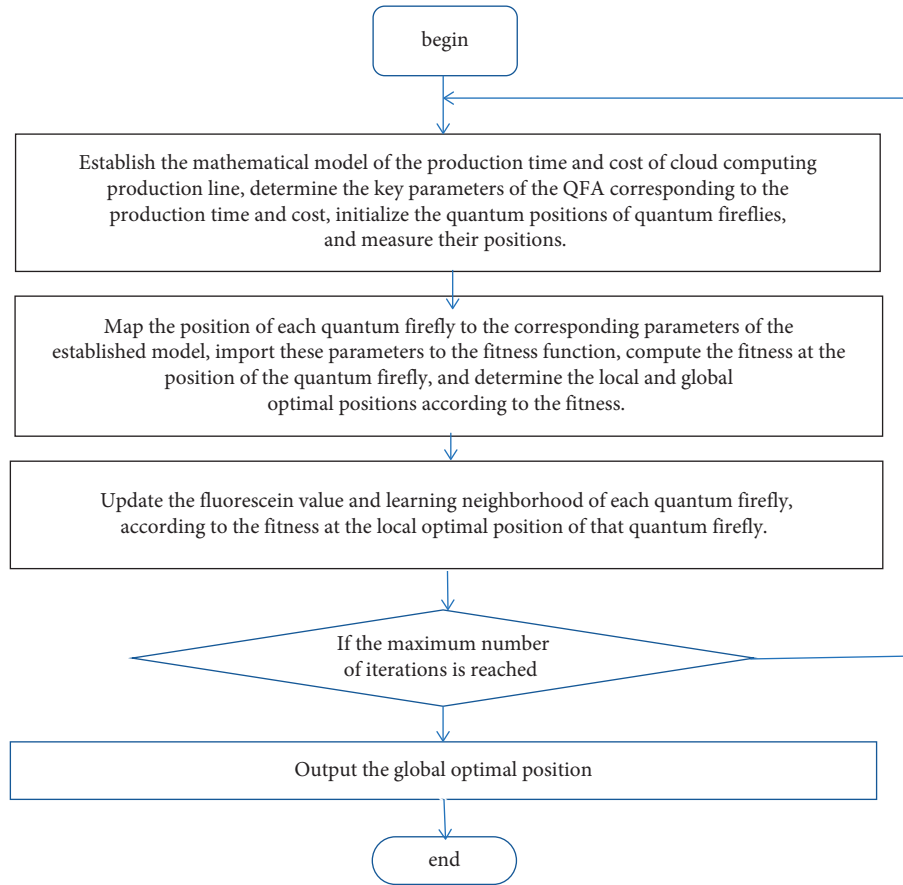


FIGURE 1: Program flow chart.

TABLE 1: Example of cloud manufacturing resource allocation.

Order	$d_i$ (10,000 yuan/1,000 pieces/day)	Suborder	Production line	$c_{ik}$ (10,000 yuan/1,000 pieces)	$l_{ik}$ /day
J <sub>1</sub>	1	$F_{1,1}$	$K_3/ K_5/ K_7$	3/3/5	5/5/3
		$F_{1,2}$	$K_1/ K_6/ K_7$	5/8/8	8/5/5
		$F_{1,3}$	$K_3/ K_5$	6/3	3/6
		$F_{1,4}$	$K_5/ K_9$	4/3	3/4
J <sub>2</sub>	1	$F_{2,1}$	$K_4/ K_9$	10/4	4/10
		$F_{2,2}$	$K_2/ K_5/ K_{10}$	8/5/4	4/5/8
		$F_{2,3}$	$K_3/ K_8$	3/5	5/3
		$F_{2,4}$	$K_1/ K_3/ K_7$	6/2/6	2/6/2
J <sub>3</sub>	1	$F_{3,1}$	$K_3/K_9/K_{10}$	4/3/4	3/4/3
		$F_{3,2}$	$K_1/K_5/K_{10}$	3/4/3	4/3/4
		$F_{3,3}$	$K_3/K_4/K_9$	3/6/4	6/3/4
J <sub>4</sub>	1	$F_{4,1}$	$K_2/K_4/K_7$	3/6/4	6/3/4
		$F_{4,2}$	$K_1/K_6/K_8$	6/8/7	8/6/7
		$F_{4,3}$	$K_1/K_4/K_{10}$	5/8/7	8/5/7
		$F_{4,4}$	$K_5/K_6/K_9$	5/9/3	5/3/9
J <sub>5</sub>	1	$F_{5,1}$	$K_5/K_6/K_9$	2/6/2	6/2/6
		$F_{5,2}$	$K_3/K_4$	3/7	7/3
		$F_{5,3}$	$K_2/K_7/K_{10}$	5/4/3	3/4/5
		$F_{5,4}$	$K_1/K_3/K_5/K_8$	4/3/4/6	4/6/4/3
J <sub>6</sub>	1	$F_{6,1}$	$K_2/K_6$	5/6	6/5
		$F_{6,2}$	$K_4/K_5/K_{10}$	6/4/6	4/6/4
		$F_{6,3}$	$K_2/K_6/K_9$	8/10/6	8/6/10
		$F_{6,4}$	$K_4/K_7/K_{10}$	10/7/5	5/7/10

TABLE 1: Continued.

Order	$d_i$ (10,000 yuan/1,000 pieces/day)	Suborder	Production line	$c_{ik}$ (10,000 yuan/1,000 pieces)	$l_{ik}$ /day
$J_7$	1	$F_{7,1}$	K3/K6	3/8	8/3
		$F_{7,2}$	K2/K8	5/5	5/5
		$F_{7,3}$	K3/K8	4/7	7/4
		$F_{7,4}$	K2/K6/K9	3/7/2	3/2/7
		$F_{7,5}$	K2/K4/K8	5/4/6	5/6/4
$J_8$	1	$F_{8,1}$	K1/K6/K10	5/6/6	6/5/5
		$F_{8,2}$	K1/K5/K7	3/3/9	9/9/3
		$F_{8,3}$	K3/K6/K9	3/6/4	6/3/4
		$F_{8,4}$	K2/K3/K8	5/3/4	3/5/4
		$F_{8,5}$	K1/K5/K8	5/3/4	3/5/4
		$F_{8,6}$	K1/K9/K10	5/4/4	5/5/4

TABLE 2: Transport cost.

Production line transport cost (10,000 yuan/1,000 pieces)/(day/1,000 pieces)	$K_1$	$K_2$	$K_3$	$K_4$	$K_5$	$K_6$	$K_7$	$K_8$	$K_9$	$K_{10}$
$K_1$	0/0	2/1	1/0.5	2/1	3/1.5	2/1	4/2	4/2	1/0.5	2/1
$K_2$	3/0.5	0/0	2/1	3/1.5	1/0.5	2/1	2/1	3/1.5	2/1	1/0.5
$K_3$	2/0.5	3/1.5	0/0	1/0.5	2/2	3/1.5	2/1	2/1	3/1.5	3/1.5
$K_4$	1/0.5	2/1	2/1	0/0	1/0.5	3/1.5	3/1.5	2/1	3/1.5	2/1
$K_5$	1/0.5	3/1.5	1/0.5	3/1.5	0/0	2/1	4/2	3/1.5	2/1	3/1.5
$K_6$	2/0.5	2/1	2/1	4/2	2/1	0/0	3/1.5	3/1.5	4/2	2/1
$K_7$	3/0.5	3/1.5	3/1.5	3/1.5	2/1	3/1.5	0/0	2/1	3/1.5	2/1
$K_8$	2/0.5	3/1.5	3/1.5	2/1	2/1	2/1	2/1	0/0	2/1	3/1.5
$K_9$	3/0.5	2/1	4/2	2/1	3/1.5	4/2	3/1.5	1/0.5	0/0	3/1.5
$R_{10}$	1/0.5	2/1	2/1	3/1.5	2/1	5/2.5	2/1	3/1.5	3/1.5	0/0

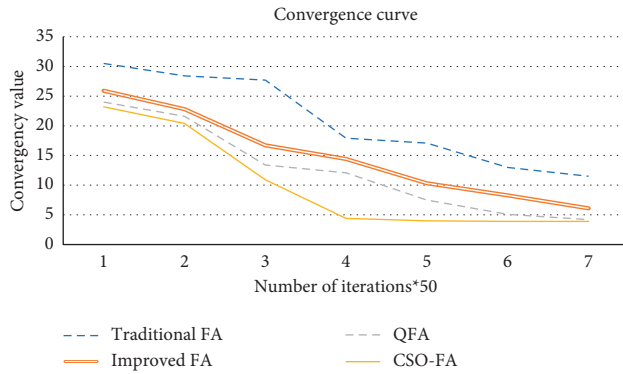


FIGURE 2: Convergence curves of different algorithms.

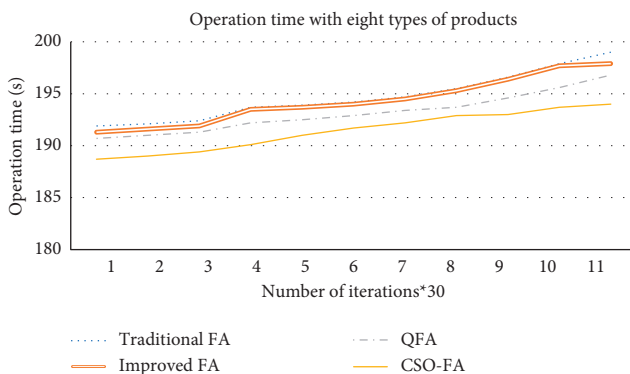


FIGURE 3: Operation time of different algorithms.

and global optimal position on quantum rotation angle,  $\theta_{il}^{t+1} = e_3(p_{il}^t - x_{il}^t) + e_4(p_{zl}^t - x_{il}^t) + e_5(p_{gl}^t - x_{il}^t)$ ;  $i = 1, 2, \dots, h$ ;  $l = 1, 2, \dots, D$ ;  $p_{zl}^t$  is the  $l$ -th dimension of the local optimal position, i.e., the position with the highest fluorescein, in the learning neighborhood of the  $i$ -th quantum firefly;  $e_3$ ,  $e_4$ , and  $e_5$  are the degree of influence of the local optimal position of the  $i$ -th quantum firefly) and  $c_1 = c_2 = 0.1/D$ .

## 5. Conclusions

This paper mainly solves the cloud robot scheduling of a cell-line involving various small batches of products, which are produced in multiple cycles, in the light of transport cost. The QFA was adopted to reasonably allocate the resources in the cloud environment. Besides, the ideas of quantum computing were added to the FA to speed up the search, while maintaining the swarm search ability. Compared with the traditional FA, improved FA, and CSO-FA, the proposed QFA (as Figures 2 and 3) achieves excellent optimization ability and a fast search speed. Finally, our algorithm was demonstrated through an example analysis on the production order scheduling of an electronic product manufacturer. The results show that the QFA is capable of effectively solving the multiobjective scheduling of cloud manufacturing resources [26, 27].

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported by Anyang Key Science and Technology Project (anke[2020]14–48; 2021A01SF007), Key Science and Technology Project, Anyang Institute of Technology, and Key Scientific Research Projects of Universities (YPY2020009), and Henan Provincial Department of Education (20B460002).

## References

- [1] Y. Tian, D. Li, P. Zhou, and L. Wang, “Coordinated scheduling of intercell production and intercell transportation in the equipment manufacturing industry,” *Engineering Optimization*, vol. 48, no. 12, pp. 2046–2063, 2016.
- [2] P. T. Zacharia and A. C. Nearchou, “A population-based algorithm for the bi-objective assembly line worker assignment and balancing problem,” *Engineering Applications of Artificial Intelligence*, vol. 49, pp. 1–9, 2016.
- [3] C. Liu, J. Lian, Y. Yin, and W. Li, “Seru Seisan—an innovation of the production management mode in Japan,” *Asian Journal of Technology Innovation*, vol. 18, no. 2, pp. 89–113, 2010.
- [4] D. I. Miyake, “The shift from belt conveyor line to work-cell based assembly systems to cope with increasing demand variation in Japanese industries,” *International Journal of Automotive Technology and Management*, vol. 6, no. 4, pp. 419–439, 2006.
- [5] Y. Du, Q. X. Chen, and J. W. Chen, “Investigation and implementation of environmental protection and energy saving,” *Advanced Materials Research*, vol. 347, pp. 3271–3274, 2012.
- [6] Q. D. Wu, Y. M. Ma, L. Li, and F. Qiao, “Data-driven dynamic scheduling method for semiconductor production line,” *Control Theory & Applications*, vol. 32, no. 9, pp. 1233–1239, 2015.
- [7] X. Cao, H. Bo, Y. Liu, and X. Liu, “Effects of different resource-sharing strategies in cloud manufacturing: A Stackelberg game-based approach,” *International Journal of Production Research*, pp. 1–21, 2021.
- [8] D. Carlucci, P. Renna, S. Materi, and G. Schiuma, “Intelligent decision-making model based on minority game for resource allocation in cloud manufacturing,” *Management Decision*, vol. 58, no. 11, pp. 2305–2325, 2020.
- [9] Y. Sakazume, “Is Japanese cell manufacturing a new system? A comparative study between Japanese cell manufacturing and cellular manufacturing (< special English Issue> production and logistics),” *Journal of Japan Industrial Management Association*, vol. 55, no. 6, pp. 341–349, 2005.
- [10] B. H. Li, L. Zhang, S. L. Wang et al., “Cloud manufacturing: a new service-oriented networked manufacturing model,” *Computer Integrated Manufacturing Systems*, vol. 16, no. 1, pp. 1–7, 2010.
- [11] Y. Du, “The research of multi-stage variable batch production line reconstruction,” *Guangdong University of Technology*, vol. 2017, no. 3, p. 127, 2017.
- [12] Y. Du and Q. X. Chen, “The model and simulation of SMT production,” *Applied Mechanics and Materials*, vol. 217–219, pp. 1493–1496, 2012.
- [13] Y. Du, F. Shi, Q. Chen, Y. Wang, J. Zhao, and Q. Li, “An improved particle swarm scheduling algorithm based on batch changing production time,” *European Journal of Electrical Engineering*, vol. 20, no. 4, pp. 439–453, 2018.
- [14] Y. Du, F. Shi, and X. Lu, “Complex robot assembly production line of new energy component,” in *Proceedings of the 2018 2nd International Conference on Artificial Intelligence: Technologies and Applications (ICAITA 2018)*, pp. 108–110, Chengdu, China, March 2018.
- [15] S. L. Wang, W. Y. Song, L. Kang, Q. Li, L. Guo, and G. S. Chen, “Manufacturing resource allocation based on cloud manufacturing,” *Computer Integrated Manufacturing Systems*, vol. 18, no. 7, pp. 1396–1405, 2012.
- [16] W. N. Liu, B. Liu, and D. H. Sun, “Multi-task oriented service composition in cloud manufacturing,” *Computer Integrated Manufacturing Systems*, vol. 19, no. 1, pp. 199–209, 2013.
- [17] C. Yin, Y. Zhang, and T. Zhong, “Optimization model of cloud manufacturing services resource combination for new product development,” *Computer Integrated Manufacturing Systems*, vol. 18, no. 7, pp. 1368–1378, 2012.
- [18] P. Kshirsagar, D. Jiang, and Z. Zhang, “Implementation and evaluation of online system identification of electromechanical systems using adaptive filters,” *IEEE Transactions on Industry Applications*, vol. 52, no. 3, pp. 2306–2314, 2016.
- [19] J. Puttonen, A. Lobov, M. A. Cavia Soto, and J. L. Martinez Lastra, “Planning-based semantic web service composition in factory automation,” *Advanced Engineering Informatics*, vol. 29, no. 4, pp. 1041–1054, 2015.
- [20] Y. Cao, Z. Wu, Q. Zhao, H. Yan, and J. Yang, “Study on resource configuration on cloud manufacturing,” *Mathematical Problems in Engineering*, vol. 2015, Article ID 635602, 2015.
- [21] Y. Cao, S. Wang, L. Kang, and Y. Gao, “A TQCS-based service selection and scheduling strategy in cloud manufacturing,” *International Journal of Advanced Manufacturing Technology*, vol. 82, no. 1–4, pp. 235–251, 2016.
- [22] Y. H. Xiong, J. Wang, M. Wu, and J. H. She, “Virtual resource scheduling method of cloud manufacturing oriented to multi-objective optimization,” *Computer Integrated Manufacturing Systems*, vol. 21, no. 11, pp. 3079–3087, 2015.
- [23] P. Ge, X. Yang, X. H. Xiao et al., “Task allocation in cloud manufacturing based on multi-resolution clustering,” *Computer Integrated Manufacturing Systems*, vol. 18, no. 7, pp. 1461–1468, 2012.
- [24] S. P. Wu, P. Zhang, and F. Li, “Multi-task scheduling based on particle swarm optimization in cloud manufacturing systems,” *Journal of South China University of Technology*, vol. 43, no. 1, pp. 105–110, 2015.
- [25] Y. Du, N. Mao, and Q. X. Chen, “Reconstruction research of the die & mould manufacturing assembly line,” *Die & Mould Industry*, vol. 42, no. 2, pp. 14–18, 2016.
- [26] S. Cheng, B. X. Zhang, M. Cao, Y. Fu, and S. Q. Hu, “Lot sizing decision in multi-period under demand uncertainty,” *Mathematics in Practice and Theory*, vol. 45, no. 11, pp. 39–43, 2015.
- [27] Y. D. Han, S. F. Dong, and B. C. Tan, “Multi-objective optimization for mixed-model assembly line balancing problem based on improved genetic algorithm,” *Computer Integrated Manufacturing Systems*, vol. 21, no. 6, pp. 1476–1485, 2015.

## Research Article

# A Novel Approach for Characterizing Solutions of Rough Optimization Problems Based on Boundary Region

Hamiden Abd El- Wahed Khalifa <sup>1,2</sup>, Dragan Pamucar <sup>3</sup>, Amina Hadj Kacem,<sup>4</sup>  
and W. A. Afifi<sup>4,5</sup>

<sup>1</sup>Department of Operations Research, Faculty of Graduate Studies for Statistical Research, Cairo University, Giza 12613, Egypt

<sup>2</sup>Department of Mathematics, College of Science and Arts, Qassim University, Al- Badaya, Saudi Arabia

<sup>3</sup>Department of Logistics, University of Defence in Belgrade, Belgrade 192204, Serbia

<sup>4</sup>Mathematics and Statistics Department, College of Science, Taibah University, Yanbu, Saudi Arabia

<sup>5</sup>Department of Mathematics, Faculty of Science, Tanta University, Tanta, Egypt

Correspondence should be addressed to Dragan Pamucar; dpamucar@gmail.com

Received 28 January 2022; Accepted 8 February 2022; Published 24 March 2022

Academic Editor: Daqing Gong

Copyright © 2022 Hamiden Abd El- Wahed Khalifa et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Rough set theory, presented by Pawlak in 1981, is one of the most well-known methods for communicating ambiguity by estimating an item based on some knowledge rather than membership. The concept of a rough function and its convexity and differentiability in regard to its boundary region are discussed in this work. The boundary notion is also used to present a new form of rough programming issue and its solutions. Finally, numerical examples are provided to demonstrate the proposed method and emphasize its advantages over other approaches.

## 1. Introduction

Rough set theory has been used to a wide range of problems. In rough set theory, knowledge is said to be dependent on the ability to classify objects, and the indiscernibility relation, which is an equivalence relation, is used to represent it formally [1]. The indiscernibility relation generates an approximation space made up of indiscernible item equivalence classes that spans the entire universe. Pawlak et al. [2] established the concept of a rough set, and one of the most prominent theories to explain ambiguity using the boundary area of a set rather than membership is Pawlak's theory [3]. A rough set, on the one hand, is distinct from ordinary and fuzzy sets in terms of concept. A characteristic function identifies an object in an ordinary set; however, in a fuzzy set, the data's uncertainty is reflected by a partial degree of membership between 0 and 1 [4]. A rough set, on the other

hand, approximates an object based on some prior knowledge. The following are some examples of rough mathematical programming problems:

- (i) *1st class*: using a rough feasible set and a crisp objective function and solving mathematical programming problems
- (ii) *2nd class*: problems in mathematical programming with a crisp feasible set and a rough target function
- (iii) *3rd class*: problems requiring a rough feasible set and a rough objective function in mathematical programming

In rough mathematical programming problems, the ideal solution set is characterized in a rough sense by four optimal sets, each spanning a distinct level of feasibility and optimality [5] (Table 1).



TABLE 1: Optimal solution set covering the different possible degree of feasibility and optimality.

	Optimality		
		Possibly	Surely
Feasibility	Possibly Surely	FOpp FOsp	FOPs FOss

- (1) FOss is the set of solutions that are certain to be feasible and certain to be optimal
- (2) FOsp signifies the set of all possible optimal solutions that are certain to be practicable
- (3) FOPs is a term that refers to a collection of all potentially viable and unquestionably optimal alternatives
- (4) FOpp signifies the collection of all possible and optimal solutions

The viable region in the second class is a crisp set; therefore, FOPs = FOss and FOpp = FOsp.

It is clear that FOss  $\subseteq$  FOsp  $\subset$  FOpp, FOss  $\subset$  FOPs  $\subset$  FOpp, and FOss = FOsp  $\cap$  FOPs.

Pawlak et al. [2] and Pawlak [3] were the first to propose the concept of a rough set [3]. Only in these instances, rough set theory is used to represent unclear data, and we just contribute to the “postprocessing step” of the data mining process. Rough multiple objective programming (RMOP) problems are the name for these innovative tasks, and they are grouped into three groups based on where the problem’s roughness appears. There are many applications for the rough set theory such as artificial intelligence, expert systems, civil engineering [6–10], medical data analysis [11], data mining (Munakata; [12, 13], Pattern recognition [14]; and [15], and decision theory [16] and [15,17–20], and [21–23]. After turning the random rough variables in the constraint set into crisp ones, Xu and Yao [24] suggested an interactive technique to solve a class of multiobjective programming problems with random rough coefficients. Osman et al. [25] investigated a method for solving a multiobjective transportation issue with rough parameters using a solution approach. Attaya [26] described and solved various objective programming problems with a degree of vagueness in their formulation. Brikaa et al. [27] solved constrained matrix games with fuzzy rough payoff matrices using an effective fuzzy multiobjective programming method. In their proposed model, Ghosh and Roy [28] built a multiobjective product mixing fixed-charge transportation problem with truckload constraints, and an extra cost that was considered as a type-I fixed charge was explored, as well as truck load limits. In a neutrosophic context, Ahmad et al. [29] proposed a new approach for addressing multilevel linear fractional programming problems, with the objective function coefficients represented by rough intervals.

The concept of a rough function, and its convexity and differentiability depending on its boundary region, which are all important concepts to understand, is introduced in this study. In addition, using the concept of a border

region, a novel sort of rough programming challenge is investigated, as well as its answers. Many authors studied the roughness in the optimization problems [30–34].

In terms of its boundary region, this research explores the concept of a rough function, as well as its convexity and differentiability, using inspiration from the above literature. Moreover, the boundary notion is also used to present a new form of rough programming issue and its solutions. It has the distinction due to the inclusion of the following feature time in literature:

- (1) Rough multiobjective programming problem
- (2) Kuhn–Tucker. Saddle point of rough programming (RP) problem
- (3) Optimal solution-based scenario

The following are the study’s key goals:

- (i) To distinguish between many forms of optimal solutions for a rough multiobjective programming issue.
- (ii) To use a numerical example to validate the suggested study

The rest of the paper can be summarized as shown in Figure 1.

## 2. Preliminaries

Some rough function definitions and convexity based on its boundary region are recalled in this part.

*Definition 1* (see [5]). In the rough mathematical programming problem, the optimum value of the objective function is a rough number  $f^*$  specified by lower and upper approximation bounds, denoted by  $f^{*(UAI)}$  and  $f^{*(LAI)}$ ; respectively.

If  $f^{*(UAI)} = f^{*(LAI)}$ , then the optimal value  $f^*$  is exact, otherwise,  $f^*$  is rough.

Roughness can be found everywhere in the rough mathematical programming problem. Rough feasibility and rough optimality are two novel concepts that have piqued our interest. Only in the first and third classes, where the feasible set is a rough set, rough feasibility does arise. The following solutions have varying degrees of feasibility:

- (i) If a solution  $x \in X$  belongs to the lower approximation of the feasible set, it is certain to be feasible
- (ii) If a solution  $x \in X$  belongs to the upper approximation of the feasible set, it is possibly feasible
- (iii) If a solution  $x \in X$  does not belong to the upper approximation of the feasible set, it is most likely not feasible
- (iv) Rough optimality can be found in a variety of rough mathematical programming problems, with variable degrees of optimality, as demonstrated below:
- (v) if  $f(x) = f^{*(UAI)}$ , the solution  $x \in X$  is unquestionably optimal
- (vi) if  $f(x) \geq f^{*(UAI)}$ , a solution is possibly optimal

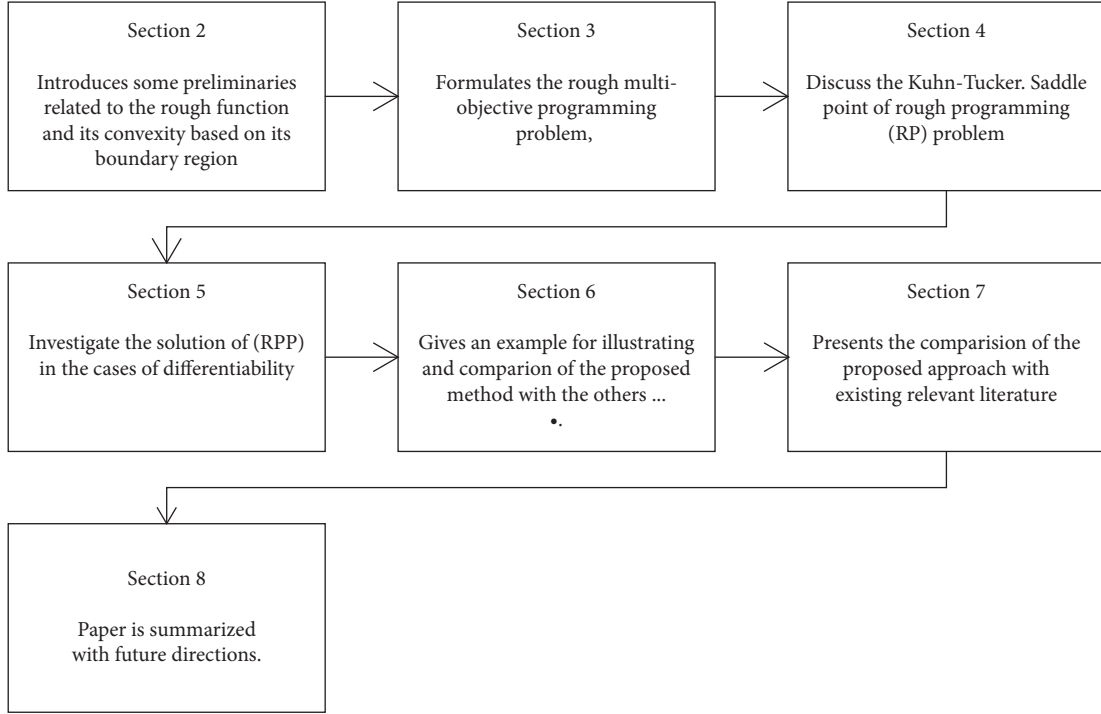


FIGURE 1: Layout of the remaining paper.

(vii) if  $(x) < f^{*(\text{UAI})}$ , a solution  $x \in X$  is unquestionably not optimal

In rough mathematical programming problems, the optimal solution sets are defined in a rough sense by four optimal sets covering the different possible degree of feasibility and optimality.

*Definition 2.* Let  $\tilde{f}^R: \mathbb{R}^n \rightarrow \mathbb{R}$  and  $r, \check{r}, r < \check{r}$ ; we suppose that a set of functions  $U (U = \{f(x), f(x): \mathbb{R}^n \rightarrow \mathbb{R}\})$  is called the universe set. The set of functions  $\{f_j\} \subset U$  is a lower approximation  $L(\tilde{f}^R(x))$  of  $\tilde{f}^R(x)$  which is denoted by  $f^{LAI}(x)$  and is defined by  $f^{LAI}(x) = \{f_j(x) \in U: |f_j - \tilde{f}^R| < r\}$ , and the set of functions  $\{f_k\} \subset U$  is an upper approximation  $U(\tilde{f}^R(x))$  of  $\tilde{f}^R(x)$  which is symbolized by  $f^{UAI}(x)$  and is characterized with  $f^{UAI}(x) = \{f_k(x) \in U: |f_k - \tilde{f}^R| < \bar{r}\}$ , where  $\{f^{LAI}(x)\} \subset \{f^{UAI}(x)\}$ . The function  $\tilde{f}^R(x)$  is called rough function if  $f^{LAI}(x) \neq f^{UAI}(x)$ .

*Definition 3.* The boundary of the rough function  $\tilde{f}^R(x)$  is  $F(x) = f^{UAI}(x) - f^{LAI}(x)$ , where  $f^{LAI}(x)$  and  $f^{UAI}(x)$  are the lower and upper approximation of  $\tilde{f}^R(x)$ .

*Definition 4.* Let  $\tilde{f}^R: \mathbb{R}^n \rightarrow \mathbb{R}$  and  $u, \hat{u} \in \mathbb{R}, u < \hat{u}$ . We suppose that the universal set  $V (V = \{f(x): tfn: q\mathbb{R}^n h \rightarrow \mathbb{R}\})$ . The set of functions  $\{f_j\} \subset V$  is the lower and upper approximation of  $\tilde{f}^R$  which is denoted by  $f^{LAI}(x)$  and  $f^{UAI}(x)$ , respectively, and they are defined as

$$f^{LAI}(x) = \left\{ f \in V: |f_j - \tilde{f}^R| < u \right\}, \quad (1)$$

$$f^{UAI}(x) = \left\{ f \in V: |f_j - \tilde{f}^R| < \hat{u} \right\}. \quad (2)$$

The function  $\tilde{f}^R$  is called rough function if

$$f^{LAI}(x) \neq f^{UAI}(x). \quad (3)$$

*Definition 5.* The boundary function of the rough function  $\tilde{f}^R$  is  $F(x) = f^{UAI}(x) - f^{LAI}(x)$ , where  $f^{LAI}$  and  $f^{UAI}$  are defined in (1) and (2), respectively.

*Definition 6.* A rough function  $\tilde{f}^R$  is said to be convex if the boundary function  $F(x)$  is convex.

*Definition 7.* Let  $X$  be an open set on  $X$ . An interval-valued function  $f: X \rightarrow \mathbb{R}$  with  $f(x) = [f^{LAI}(x), f^{UAI}(x)]$  is called weakly differentiable at  $x_0 \in X$  if the real-valued functions  $f^{LAI}$  and  $f^{UAI}$  are differentiable at  $x_0$ .

### 3. Problem Statement and Solution Concepts

A rough programming (RP) problem can be stated as

$$(RP) \min \tilde{f}^R(x). \quad (4)$$

Subject to

$$X = \{x \in \mathbb{R}^n: h_r(x) \leq 0, r = \overline{1, m}\}, \quad (5)$$

where  $f_k^{LAI}(x)$  and  $f_k^{UAI}(x)$  are the lower and upper approximations of  $\tilde{f}^R(x)$ , respectively, and  $f_k^{LAI}(x) \leq$

$\tilde{f}_k^R(x) \leq f_k^{\text{UAI}}(x)$ ,  $k = \overline{1, \overline{K}}$ , and  $X$  represents the convex crisp feasible region, and  $h_r(x)$ ,  $r = 1, 2, \dots, m$  are the convex and continuous functions.

In order to solve the (RP) problem, let us solve the following boundary problem (BP):

$$(BP) \min F(x) = f^{\text{UAI}}(x) - f^{\text{LAI}}(x). \quad (6)$$

Subject to

$$X = \{x \in \mathbb{R}^n : h_r(x) \leq 0, r = 1, 2, \dots, m\}, \quad (7)$$

where  $X$  is the convex set and  $h_r(x)$ ,  $r = 1, 2, \dots, m$  are the convex and continuous functions.

The BP can be separated into the following two sub-problems as

$$(LAP) \max f^{\text{LAI}}(x). \quad (8)$$

Subject to

$$X = \{x \in \mathbb{R}^n : h_r(x) \leq 0, r = \overline{1, m}\}, \quad (9)$$

$$(UAP) \min f^{\text{UAI}}(x).$$

Subject to

$$X = \{x \in \mathbb{R}^n : h_r(x) \leq 0, r = 1, 2, \dots, m\}, \quad (10)$$

where  $f^{\text{LAI}}(x)$  and  $f^{\text{UAI}}(x)$  are the concave and convex functions, respectively.

Let the optimal solutions of (LAP) and (UAP) be denoted by  $f^{\text{LAI}}(x^*) = \max_{x \in X} f^{\text{LAI}}(x)$ , and

$$f^{\text{UAI}}(x^*) = \min_{x \in X} f^{\text{UAI}}(x), \quad (11)$$

respectively.

**Definition 8.**  $x^*$  is said to be the optimal solution of the RP problem if  $f^{\text{LAI}}(x^*) \leq \tilde{f}^R(x^*) \leq f^{\text{UAI}}(x^*)$  where  $S^L$  and  $S^U$  are the sets of the solutions of problems (LAP) and (UAP), respectively.

**Definition 9**

- (1) A solution  $x^* \in S^L \cap S^U$ ,  $F(x^*) = 0$  is called a surely optimal solution of the RP
- (2)  $x^* \in S^L \cap S^U$ ,  $F(x^*) \neq 0$  is called a possibly optimal solution of the RP
- (3)  $x^* \in S^L \cap S^U$  is called a nearly possibly optimal solution of the RP

**Lemma 1.** If  $x^*$  is the solution of (BP), then  $x^*$  is the solution for (LAP) and (UAP).

*Proof.* Let  $x^*$  be a solution of BP; then,

$$f^{\text{UAI}}(x^*) - f^{\text{LAI}}(x^*) \leq f^{\text{UAI}}(x) - f^{\text{LAI}}(x); \forall x. \quad (12)$$

We suppose that  $x^*$  is not a solution for (LAP) and (UAP), then there exists an  $\bar{x} \in X$  such that  $f^{\text{UAI}}(\bar{x}) \leq f^{\text{UAI}}(x^*)$  implies that

$$f^{\text{UAI}}(\bar{x}) - f^{\text{LAI}}(\bar{x}) < f^{\text{UAI}}(x^*) - f^{\text{LAI}}(\bar{x}), f^{\text{LAI}}(x^*) < f^{\text{LAI}}(\bar{x}) \text{ which leads to}$$

$$f^{\text{UAI}}(x^*) - f^{\text{LAI}}(x^*) > f^{\text{UAI}}(x^*) - f^{\text{LAI}}(\bar{x}). \quad (13)$$

Thus,  $f^{\text{UAI}}(\bar{x}) - f^{\text{LAI}}(\bar{x}) < f^{\text{UAI}}(x^*) - f^{\text{LAI}}(x^*)$  contradicts that  $x^*$  is a solution of BP. Therefore,  $x^*$  is a solution of the two problems (LAP) and (UAP).  $\square$

#### 4. Rough Kuhn–Tucker Saddle Point

We consider the rough problem

$$\min \tilde{f}^R(x). \quad (14)$$

Subject to

$$X = \{x \in \mathbb{R}^n : h_r(x) \leq 0, r = \overline{1, m}\}, \quad (15)$$

$$f^{\text{LAI}}(x) \leq \tilde{f}^R(x) \leq f^{\text{UAI}}(x).$$

The rough Kuhn–Tucker saddle point for problem (15) takes the form

$$\begin{aligned} & \tilde{f}^R(x^*) + \sum_{r=1}^m \gamma_r h_r(x^*) + \gamma_{m+1} (f^{\text{LAI}}(x^*) - \tilde{f}^R(x^*)) + \gamma_{m+2} (\tilde{f}^R(x^*) - f^{\text{UAI}}(x^*)), \\ & \leq \tilde{f}^R(x^*) + \sum_{r=1}^m \gamma_r^* h_r(x^*) + \gamma_{m+1}^* (f^{\text{LAI}}(x^*) - \tilde{f}^R(x^*)) + \gamma_{m+2}^* (\tilde{f}^R(x^*) - f^{\text{UAI}}(x^*)), \\ & \leq \tilde{f}^R(x) + \sum_{r=1}^m \gamma_r^* h_r(x) + \gamma_{m+1}^* (f^{\text{LAI}}(x) - \tilde{f}^R(x)) + \gamma_{m+2}^* (\tilde{f}^R(x) - f^{\text{UAI}}(x)), \end{aligned} \quad (16)$$

or

$$\begin{aligned}
& (1 - \gamma_{m+1} + \gamma_{m+2})\tilde{f}^R(x^*) + \sum_{r=1}^m \gamma_r h_r(x^*) + \gamma_{m+1} f^{\text{LAI}}(x^*) - \gamma_{m+2} f^{\text{UAI}}(x^*), \\
& \leq (1 - \gamma_{m+1}^* + \gamma_{m+2}^*)\tilde{f}^R(x^*) + \sum_{r=1}^m \gamma_r^* h_r(x^*) + \gamma_{m+1}^* f^{\text{LAI}}(x^*) - \gamma_{m+2}^* f^{\text{UAI}}(x^*), \\
& \leq (1 - \gamma_{m+1}^* + \gamma_{m+2}^*)\tilde{f}^R(x) + \sum_{r=1}^m \gamma_r^* h_r(x) + \gamma_{m+1}^* f^{\text{LAI}}(x) - \gamma_{m+2}^* f^{\text{UAI}}(x).
\end{aligned} \tag{17}$$

**Theorem 1.** If  $(x^*, \gamma_r^*)$ , where  $\gamma_r^* \geq 0, r = \overline{1, m+2}$ , and  $\sum_{r=1}^{m+1} \gamma_r^*$  is a rough Kuhn-Tucker saddle point (KTSP), then  $x^*$  is a solution of RP.

*Proof.* We assume that  $(x^*, \gamma_r^*), r = \overline{1, m+2}$  is a rough KTSP; then, for  $\gamma_r \geq 0, \gamma_r \in \mathbb{R}^{m+2}$ , we get

$$\begin{aligned}
& (1 - \gamma_{m+1} + \gamma_{m+2})\tilde{f}^R(x^*) + \sum_{r=1}^m \gamma_r h_r(x^*) + \gamma_{m+1} f^{\text{LAI}}(x^*) - \gamma_{m+2} f^{\text{UAI}}(x^*), \\
& \leq (1 - \gamma_{m+1}^* + \gamma_{m+2}^*)\tilde{f}^R(x^*) + \sum_{r=1}^m \gamma_r^* h_r(x^*) + \gamma_{m+1}^* f^{\text{LAI}}(x^*) - \gamma_{m+2}^* f^{\text{UAI}}(x^*), \\
& \leq (1 - \gamma_{m+1}^* + \gamma_{m+2}^*)\tilde{f}^R(x) + \sum_{r=1}^m \gamma_r^* h_r(x) + \gamma_{m+1}^* f^{\text{LAI}}(x) - \gamma_{m+2}^* f^{\text{UAI}}(x).
\end{aligned} \tag{18}$$

From the first inequality, we have

$$\begin{aligned}
& (1 - \gamma_{m+1} + \gamma_{m+2})\tilde{f}^R(x^*) + \sum_{r=1}^m \gamma_r h_r(x^*) + \gamma_{m+1} f^{\text{LAI}}(x^*) - \gamma_{m+2} f^{\text{UAI}}(x^*), \\
& \leq (1 - \gamma_{m+1}^* + \gamma_{m+2}^*)\tilde{f}^R(x) + \sum_{r=1}^m \gamma_r^* h_r(x) + \gamma_{m+1}^* f^{\text{LAI}}(x) - \gamma_{m+2}^* f^{\text{UAI}}(x),
\end{aligned} \tag{19}$$

or

$$\begin{aligned}
& (1 - \gamma_{m+1} + \gamma_{m+2} + 1 - \gamma_{m+1}^* + \gamma_{m+2}^*)\tilde{f}^R(x^*) + \sum_{r=1}^m (\gamma_r - \gamma_r^*) h_r(x^*) \\
& + (\gamma_{m+1} - \gamma_{m+1}^*) f^{\text{LAI}}(x^*) - (\gamma_{m+2} - \gamma_{m+2}^*) f^{\text{UAI}}(x^*) \leq 0,
\end{aligned} \tag{20}$$

which implies to

$$\left( \gamma_{m+1} - \gamma_{m+1}^* \right) \left( f^{\text{LAI}}(x^*) - \tilde{f}^R(x^*) \right) + \left( \gamma_{m+2} - \gamma_{m+2}^* \right) \left( \tilde{f}^R(x^*) - f^{\text{UAI}}(x^*) \right) + \sum_{r=1}^m (\gamma_r - \gamma_r^*) h_r(x^*) \leq 0. \tag{21}$$

This inequality is true for all  $\gamma_r, \gamma_r^*, \gamma_{m+1}, \gamma_{m+1}^*, \gamma_{m+2}, \gamma_{m+2}^*$ . In the case  $\gamma_{m+1} = \gamma_{m+1}^*$  and  $\gamma_{m+2} = \gamma_{m+2}^*$ , we have  $\sum_{r=1}^m (\gamma_r - \gamma_r^*) h_r(x^*) \leq 0$ . We assume that  $\gamma_r = \gamma_r^*, r = 1, 2, \dots, i-1, i+1, \dots, m$  and  $\gamma_i^* = \gamma_i - 1$ . Then,  $h_r(x^*) \leq 0$ . By repeating this for all  $i$ , we have

$h_r(x^*) \leq 0$ , and hence,  $x^*$  is the feasible point. Since  $\gamma_r^* \geq 0$  and  $h_r(x^*) \leq 0$ , we get  $\sum_{r=1}^m \gamma_r^* h_r(x^*) \leq 0$ . Again from the first inequality, where  $\gamma_{m+1} = \gamma_{m+1}^*$  and  $\gamma_{m+2} = \gamma_{m+2}^*$ , and by setting  $\gamma_r$ , we obtain  $\sum_{r=1}^m \gamma_r^* h_r(x^*) \geq 0$ . Hence,  $\sum_{r=1}^m \gamma_r^* h_r(x^*) = 0$ . Thus,

$$(\gamma_{m+1} - \gamma_{m+1}^*)(f^{\text{LAI}}(x^*) - \tilde{f}^R(x^*)) + (\gamma_{m+2} - \gamma_{m+2}^*)(\tilde{f}^R(x^*) - f^{\text{UAI}}(x^*)) + \sum_{r=1}^m (\gamma_r - \gamma_r^*)h_r(x^*) \leq 0. \quad (22)$$

By taking  $\gamma_{m+1} = \gamma_{m+1}^* - 1$  and  $\gamma_{m+2} = \gamma_{m+2}^* - 1$ , we have  $(\gamma_{m+1} - 1 - \gamma_{m+1}^*)(f^{\text{LAI}}(x^*) - \tilde{f}^R(x^*)) + (\gamma_{m+2} - 1 - \gamma_{m+2}^*)(\tilde{f}^R(x^*) - f^{\text{UAI}}(x^*)) + \sum_{r=1}^m \gamma_r h_r(x^*) \leq 0$ . This leads to

$$\begin{aligned} & (f^{\text{LAI}}(x^*) - \tilde{f}^R(x^*)) + (\tilde{f}^R(x^*) - f^{\text{UAI}}(x^*)) \\ & + \sum_{r=1}^m \gamma_r h_r(x^*) \leq 0. \end{aligned} \quad (23)$$

Since the inequality is valid for each  $\gamma_r \geq 0$ , then for  $\gamma_r = 0$ , we get

$$\begin{aligned} & (f^{\text{LAI}}(x^*) - \tilde{f}^R(x^*)) + (\tilde{f}^R(x^*) - f^{\text{UAI}}(x^*)) \leq 0, \\ & f^{\text{UAI}}(x^*) - f^{\text{LAI}}(x^*) \leq 0. \end{aligned} \quad (24)$$

Taking  $\gamma_{m+1} = \gamma_{m+1}^* + 1$  and  $\gamma_{m+2} = \gamma_{m+2}^* + 1$ , we have

$$(\gamma_{m+1} + 1 - \gamma_{m+1}^*)(f^{\text{LAI}}(x^*) - \tilde{f}^R(x^*)) + (\gamma_{m+2} + 1 - \gamma_{m+2}^*)(\tilde{f}^R(x^*) - f^{\text{UAI}}(x^*)) + \sum_{r=1}^m \gamma_r h_r(x^*) \leq 0. \quad (25)$$

Thus,

$$\begin{aligned} & (f^{\text{LAI}}(x^*) - \tilde{f}^R(x^*)) + (\tilde{f}^R(x^*) - f^{\text{UAI}}(x^*)) \\ & + \sum_{r=1}^m \gamma_r h_r(x^*) \leq 0. \end{aligned} \quad (26)$$

Since the inequality is valid for each  $\gamma_r \geq 0$ , then for  $\gamma_r = 0$ , we have

$$\begin{aligned} & (f^{\text{LAI}}(x^*) - \tilde{f}^R(x^*)) + (\tilde{f}^R(x^*) - f^{\text{UAI}}(x^*)) \leq 0, \\ & f^{\text{UAI}}(x^*) - f^{\text{LAI}}(x^*) \geq 0. \end{aligned} \quad (27)$$

Hence from (24) and (27), we conclude  $f^{\text{LAI}}(x^*) = \tilde{f}^R(x^*) = f^{\text{UAI}}(x^*)$  (i.e.,  $x^*$  is a surely optimal solution for the RP)

From the second inequality, we have

$$\begin{aligned} & (1 - \gamma_{m+1}^* + \gamma_{m+2}^*)\tilde{f}^R(x^*) + \sum_{r=1}^m \gamma_r^* h_r(x^*) + \gamma_{m+1}^* f^{\text{LAI}}(x^*) - \gamma_{m+2}^* f^{\text{UAI}}(x^*), \\ & \leq (1 - \gamma_{m+1}^* + \gamma_{m+2}^*)\tilde{f}^R(x) + \sum_{r=1}^m \gamma_r^* h_r(x) + \gamma_{m+1}^* f^{\text{LAI}}(x) - \gamma_{m+2}^* f^{\text{UAI}}(x). \end{aligned} \quad (28)$$

Since  $\sum_{r=1}^m \gamma_r^* h_r(x^*) = 0$ , then

$$\begin{aligned} & (1 - \gamma_{m+1}^* + \gamma_{m+2}^*)(\tilde{f}^R(x^*) - \tilde{f}^R(x)) \leq \sum_{r=1}^m \gamma_r^* h_r(x) + \gamma_{m+1}^* (f^{\text{LAI}}(x) - f^{\text{LAI}}(x^*)) \\ & + \gamma_{m+2}^* (f^{\text{UAI}}(x) - f^{\text{UAI}}(x^*)) \tilde{f}^R(x^*) - \tilde{f}^R(x) \leq \frac{\sum_{r=1}^m \gamma_r^*}{(1 - \gamma_{m+1}^* + \gamma_{m+2}^*)} h_r(x) \\ & + \frac{\gamma_{m+1}^*}{(1 - \gamma_{m+1}^* + \gamma_{m+2}^*)} (f^{\text{LAI}}(x) - f^{\text{LAI}}(x^*)) + \frac{\gamma_{m+2}^*}{(1 - \gamma_{m+1}^* + \gamma_{m+2}^*)} (f^{\text{UAI}}(x) - f^{\text{UAI}}(x^*)). \end{aligned} \quad (29)$$

For  $x^* \in S^L \cap S^U$ , we have  $f^{\text{LAI}}(x) \leq f^{\text{LAI}}(x^*)$  and  $f^{\text{UAI}}(x) \geq f^{\text{UAI}}(x^*)$ . Since  $\sum_{r=1}^{m+1} \gamma_r = 1$  and  $\gamma_{m+1}^* = \gamma_1^* +$

$\gamma_1^* + \dots + \gamma_m^*$ , then  $1 - \gamma_{m+1}^* + \gamma_{m+2}^* \leq 0$  which implies to  $\tilde{f}^R(x^*) \leq \tilde{f}^R(x)$ ,  $x \in X$ . Hence,  $x^*$  is a possible

optimalsolution of the rough problem. For  $x^* \in S^L, x^* \notin S^U$ , we obtain  $f^{LAI}(x^*) \geq f^{LAI}(x)$  and

$$\tilde{f}^R(x^*) - \tilde{f}^R(x) \leq \frac{\gamma_{m+2}^*}{(1 - \gamma_{m+1}^* + \gamma_{m+2}^*)} (f^{UAI}(x) - f^{UAI}(x^*)). \quad (30)$$

Now, there are two cases:  $\square$

*Case 1.*  $f^{UAI}(x^*) - f^{UAI}(x) \leq 0; \forall x \in X$  implies that  $x^*$  is a nearly possibly optimal solution.

*Case 2.*  $f^{UAI}(x^*) - f^{UAI}(x) > 0$ .

Let  $x^*$  be not a nearly possible optimal solution of rough problem; then, there is  $\bar{x} \in X: \tilde{f}^R(\bar{x}) < \tilde{f}^R(x^*)$ . Since  $x^* \in S^L, x^* \notin S^U$ , so  $x^*$  is not a solution for the boundary problem BP, i.e., there is  $\bar{x}$ :

$$\begin{aligned} f^{UAI}(\bar{x}) - f^{LAI}(\bar{x}) &< f^{UAI}(x^*) - f^{LAI}(x^*), f^{LAI}(x^*) \\ &- f^{LAI}(\bar{x}) < f^{UAI}(x^*) - f^{UAI}(\bar{x}). \end{aligned} \quad (31)$$

(i) If  $f^{UAI}(x^*) < f^{UAI}(\bar{x})$ , then  $f^{LAI}(x^*) < f^{LAI}(\bar{x})$ . This contradicts that  $x^* \in S^L$ , and hence,  $x^*$  must be a nearly possible optimal solution for the RP problem.

(ii) If  $f^{UAI}(x^*) > f^{UAI}(\bar{x})$ , then we may write  $f^{UAI}(x^*) = f^{UAI}(\bar{x}) + \theta, \theta > 0$ , which implies to  $f^{LAI}(x^*) - f^{LAI}(\bar{x}) < \theta, \theta > 0$ . Then, we have two cases:

- (a)  $f^{LAI}(x^*) > f^{LAI}(\bar{x})$ , which is not considered, where  $x^* \in S^L$
- (b)  $f^{LAI}(x^*) < f^{LAI}(\bar{x})$ , which contradicts that  $x^* \in S^L$ , and hence,  $x^*$  must be a nearly possible optimal solution for the RP problem

For  $x^* \in S^U, x^* \notin S^L$ , we obtain  $f^{UAI}(x^*) \leq f^{UAI}(x)$  and

$$\tilde{f}^R(x^*) - \tilde{f}^R(x) \leq \frac{\gamma_{m+1}^*}{(1 - \gamma_{m+1}^* + \gamma_{m+2}^*)} (f^{LAI}(x) - f^{LAI}(x^*)). \quad (32)$$

So, there are two cases:

*Case 3.*  $f^{LAI}(x^*) - f^{LAI}(x) \leq 0; \forall x \in X$ ; this implies that  $x^*$  is a nearly possibly optimal solution.

*Case 4.*  $f^{LAI}(x^*) - f^{LAI}(x) > 0$ .

Let  $x^*$  be not a nearly possible optimal solution of the rough problem; then, there is  $\bar{x} \in X: \tilde{f}^R(\bar{x}) < \tilde{f}^R(x^*)$ . Since  $x^* \in S^U, x^* \notin S^L$ , so  $x^*$  is not a solution for the boundary problem (BP), i.e., there is  $\bar{x} \in X$ :

$$\begin{aligned} f^{UAI}(\bar{x}) - f^{LAI}(\bar{x}) &< f^{UAI}(x^*) - f^{LAI}(x^*), f^{UAI}(\bar{x}) \\ &- f^{UAI}(x^*) < f^{LAI}(\bar{x}) - f^{LAI}(x^*). \end{aligned} \quad (33)$$

(i) If  $f^{LAI}(\bar{x}) < f^{UAI}(x^*)$ , then  $f^{UAI}(\bar{x}) < f^{LAI}(x^*)$ . This contradicts that  $x^* \in S^U$ , and hence,  $x^*$  must be a nearly possible optimal solution for the (R – MOP) problem.

(ii) If  $f^{LAI}(\bar{x}) > f^{UAI}(x^*)$ , then we may write  $f^{LAI}(x^*) = f^{LAI}(\bar{x}) + \theta, \theta > 0$ , which implies to  $f^{UAI}(\bar{x}) - f^{UAI}(x^*) < \theta, \theta > 0$ . Then, we have two cases:

- (a)  $f^{LAI}(x^*) > f^{LAI}(\bar{x})$ , which is not considered, where  $x^* \in S^U$
- (b)  $f^{UAI}(x^*) < f^{UAI}(\bar{x})$ , which contradicts that  $x^* \in S^U$ , and hence,  $x^*$  must be a nearly possible optimal solution for the RP problem

## 5. Rough Function Differentiability

A rough function  $\tilde{f}^R(x)$  is said to be differentiable if its boundary.

$F(x) = f^{UAI} - f^{LAI}$  is differentiable. Then,  
 $F - F(x^*) = (\delta/\delta x)F(x^*)(x - x^*) + \vartheta(x^*, \gamma(x - x^*))$   
 $\|x - x^*\|$ , or equivalently  
 $\tilde{f}^R - \tilde{f}^R(x^*) = (\delta/\delta x)\tilde{f}^R(x^*)(x - x^*) + \vartheta(x^*, \gamma(x - x^*))$   
 $\|x - x^*\|$ , where

$$\lim_{\vartheta \rightarrow 0} \vartheta(x^*, \delta(x - x^*)) = 0. \quad (34)$$

*5.1. Kuhn–Tucker's Conditions under Roughness.* The rough Kuhn–Tucker (KT) conditions for the RP problem takes the form

$$\begin{aligned} \frac{\delta}{\delta x} \tilde{f}^R(x^*) + \sum_{r=1}^m \gamma_r^* h_r(x^*) + \gamma_{m+1}^* \frac{\delta}{\delta x} (f^{LAI}(x^*) - \tilde{f}^R(x^*)) + \gamma_{m+2}^* \frac{\delta}{\delta x} (\tilde{f}^R(x^*) - f^{UAI}(x^*)) &= 0, \\ \gamma_r^* h_r(x^*) = 0, r = \overline{1, m}; \gamma_{m+2}^* (\tilde{f}^R(x^*) - f^{UAI}(x^*)) &= 0; \gamma_r^* \geq 0, r = \overline{1, m+2}. \end{aligned} \quad (35)$$

Let  $\sum_{r=1}^{m+1} \gamma_r^* = 1$ . Then,



$$(1 - \gamma_{m+1}^* + \gamma_{m+2}^*) \frac{\delta}{\delta x} \tilde{f}^R(x^*) + \gamma_{m+1}^* \frac{\delta}{\delta x} f^{\text{LAI}}(x^*) - \gamma_{m+2}^* \frac{\delta}{\delta x} f^{\text{UAI}}(x^*) + \sum_{r=1}^m \gamma_r^* \frac{\delta}{\delta x} h_r(x^*) = 0, \quad (36)$$

or, in other words

$$\begin{aligned} & \frac{\delta}{\delta x} \tilde{f}^R(x^*) + \frac{\gamma_{m+1}^*}{(1 - \gamma_{m+1}^* + \gamma_{m+2}^*)} \frac{\delta}{\delta x} f^{\text{LAI}}(x^*) - \frac{\gamma_{m+2}^*}{(1 - \gamma_{m+1}^* + \gamma_{m+2}^*)} \frac{\delta}{\delta x} f^{\text{UAI}}(x^*) \\ & + \frac{\sum_{r=1}^m \gamma_r^*}{(1 - \gamma_{m+1}^* + \gamma_{m+2}^*)} \frac{\delta}{\delta x} h_r(x^*) = 0 \\ & \frac{\sum_{r=1}^m \gamma_r^*}{(1 - \gamma_{m+1}^* + \gamma_{m+2}^*)} \frac{\delta}{\delta x} h_r(x^*) = 0, \\ & r = \overline{1, m}, \end{aligned} \quad (37)$$

$$\frac{\gamma_{m+1}^*}{(1 - \gamma_{m+1}^* + \gamma_{m+2}^*)} f^{\text{LAI}}(x^*) = 0$$

$$\frac{\gamma_{m+2}^*}{(1 - \gamma_{m+1}^* + \gamma_{m+2}^*)} f^{\text{UAI}}(x^*)$$

$$\gamma_r^* \geq 0,$$

$$r = \overline{1, m+2}.$$

**Theorem 2.** Let  $\tilde{f}^R$ ,  $f^{\text{UAI}}$ , and  $h$  be the convex and differentiable functions at  $x^*$ , and let  $f^{\text{LAI}}$  be a concave and differentiable at  $x^* \in X$ . We suppose that  $f^{\text{UAI}}(x^*) > 0$  and  $f^{\text{LAI}}(x^*) > 0$ . If  $(x^*, \gamma_r^*)$ , where  $\gamma_r^* \geq 0, r = \overline{1, m+2}$  is a solution of the KT conditions, then  $x^*$  is a solution for RP

*Proof.* Let  $(x^*, \gamma_r^*)$  be a solution of the rough KT conditions. Since  $\tilde{f}^R$  is a convex and differentiable at  $x^*$ , we get  $\tilde{f}^R - \tilde{f}^R(x^*) \geq \delta x / \delta \tilde{f}^R(x^*)(x - x^*)$ . Since  $\delta / \delta x \tilde{f}^R(x^*) = \gamma_{m+2}^* / (1 - \gamma_{m+1}^* + \gamma_{m+2}^*) \delta / \delta x f^{\text{UAI}}(x^*) - \gamma_{m+1}^* / (1 - \gamma_{m+1}^* + \gamma_{m+2}^*) \delta / \delta x f^{\text{LAI}}(x^*) - \sum_{r=1}^m \gamma_r^* / (1 - \gamma_{m+1}^* + \gamma_{m+2}^*) h_r(x^*)$ , and  $f^{\text{UAI}}$ ,  $f^{\text{LAI}}$ , and  $h_r$  are differentiable, then

$$f^{\text{UAI}} - f^{\text{UAI}}(x^*) = \frac{\delta}{\delta x} f^{\text{UAI}}(x^*)(x - x^*) + \vartheta(x^*, \gamma(x - x^*))x - x^*,$$

$$f^{\text{LAI}} - f^{\text{LAI}}(x^*) = \frac{\delta}{\delta x} f^{\text{LAI}}(x^*)(x - x^*) + \vartheta(x^*, \gamma(x - x^*))x - x^*, \quad (38)$$

$$h_r - h_r(x^*) = \frac{\delta}{\delta x} h_r(x^*)(x - x^*) + \vartheta(x^*, \gamma(x - x^*))x - x^*.$$

Then,

$$\begin{aligned}
\tilde{f}^R - \tilde{f}^R(x^*) &\geq \frac{\gamma_{m+2}^*}{(1 - \gamma_{m+1}^* + \gamma_{m+2}^*)} (f^{UA} - f^{UA}(x^*) - \vartheta(x^*, \gamma(x - x^*))x - x^*) \\
&\quad - \frac{\gamma_{m+1}^*}{(1 - \gamma_{m+1}^* + \gamma_{m+2}^*)} (f^{LAI} - f^{LAI}(x^*) - \vartheta(x^*, \gamma(x - x^*))x - x^*) \\
&\quad - \frac{\sum_{r=1}^m \gamma_r^*}{(1 - \gamma_{m+1}^* + \gamma_{m+2}^*)} (h_r - h_r(x^*) - \vartheta(x^*, \gamma(x - x^*))x - x^*).
\end{aligned} \tag{39}$$

Since  $\lim_{\vartheta \rightarrow 0} \vartheta(x^*, \delta(x - x^*)) = 0$ , then

$$\begin{aligned}
\tilde{f}^R - \tilde{f}^R(x^*) &\geq \frac{\gamma_{m+2}^*}{(1 - \gamma_{m+1}^* + \gamma_{m+2}^*)} (f^{UAI} - f^{UAI}(x^*)) - \frac{\gamma_{m+1}^*}{(1 - \gamma_{m+1}^* + \gamma_{m+2}^*)} (f^{LAI} - f^{LAI}(x^*)) \\
&\quad - \frac{\sum_{r=1}^m \gamma_r^*}{(1 - \gamma_{m+1}^* + \gamma_{m+2}^*)} (h_r - h_r(x^*)).
\end{aligned} \tag{40}$$

From the Kuhn–Tucker conditions,

$$\begin{aligned}
\frac{\sum_{r=1}^m \gamma_r^*}{(1 - \gamma_{m+1}^* + \gamma_{m+2}^*)} h_r(x^*) &= 0, \\
r &= \overline{1, m}, \\
\frac{\gamma_{m+1}^*}{(1 - \gamma_{m+1}^* + \gamma_{m+2}^*)} f^{LAI}(x^*) &= 0, \\
\frac{\gamma_{m+2}^*}{(1 - \gamma_{m+1}^* + \gamma_{m+2}^*)} f^{UAI}(x^*) &= 0.
\end{aligned} \tag{41}$$

Then, the following inequality  $\tilde{f}^R(x) - \tilde{f}^R(x^*) \geq \frac{\gamma_{m+2}^*}{(1 - \gamma_{m+1}^* + \gamma_{m+2}^*)} f^{UAI} - \frac{\gamma_{m+1}^*}{(1 - \gamma_{m+1}^* + \gamma_{m+2}^*)} f^{LAI} - \frac{\sum_{r=1}^m \gamma_r^*}{(1 - \gamma_{m+1}^* + \gamma_{m+2}^*)} h_r$  is valid for each  $\gamma_r^* \geq 0, r = \overline{1, m+2}$ , and for  $\gamma_r^* = 0$ , we have

$$\tilde{f}^R - \tilde{f}^R(x^*) \geq \frac{\gamma_{m+2}^*}{(1 - \gamma_{m+1}^* + \gamma_{m+2}^*)} f^{UAI} - \frac{\gamma_{m+1}^*}{(1 - \gamma_{m+1}^* + \gamma_{m+2}^*)} f^{LAI}. \tag{42}$$

If  $\gamma_{m+1}^*; \gamma_{m+2}^* > 0$ , then from the Kuhn–Tucker conditions, we obtain  $\tilde{f}^R(x^*) = f^{LAI}(x^*)$  and  $\tilde{f}^R(x^*) = f^{UAI}(x^*)$ . Then,  $x^*$  is a surely optimal solution of the RP problem.

If  $x^* \in S^L \cap S^U$ , then  $f^{UAI}(x^*) \leq f^{UAI}(x); \forall x \in X$  and  $f^{LAI}(x^*) \geq f^{LAI}(x); \forall x \in X$ , and then we grt

$$\begin{aligned}
\tilde{f}^R - \tilde{f}^R(x^*) &\geq \frac{\gamma_{m+2}^*}{(1 - \gamma_{m+1}^* + \gamma_{m+2}^*)} f^{UAI}(x^*) \\
&\quad - \frac{\gamma_{m+1}^*}{(1 - \gamma_{m+1}^* + \gamma_{m+2}^*)} f^{LAI}(x^*).
\end{aligned} \tag{43}$$

In addition, from the Kuhn–Tucker conditions  $\tilde{f}^R - \tilde{f}^R(x^*) \geq 0$ , this leads to  $\tilde{f}^R(x^*) \leq \tilde{f}^R(x)$ , i.e.,  $x^*$  is a possibly optimal solution.

If  $x^* \in S^L, x^* \notin S^U$ , then  $f^{LAI}(x^*) \geq f^{LAI}(x); \forall x \in X$ , and we have

$$\begin{aligned}
\tilde{f}^R - \tilde{f}^R(x^*) &\geq \frac{\gamma_{m+2}^*}{(1 - \gamma_{m+1}^* + \gamma_{m+2}^*)} f^{UAI}(x^*) - \frac{\gamma_{m+1}^*}{(1 - \gamma_{m+1}^* + \gamma_{m+2}^*)} f^{LAI}(x^*), \\
\tilde{f}^R - \tilde{f}^R(x^*) &\geq \frac{\gamma_{m+2}^*}{(1 - \gamma_{m+1}^* + \gamma_{m+2}^*)} f^{UAI}.
\end{aligned} \tag{44}$$

From the assumption that  $f^{UAI}(x^*) > 0$ , and  $x^*$  is not solution for BP,  $\gamma_{m+2}^*/(1 - \gamma_{m+1}^* + \gamma_{m+2}^*) = 0$ .

Hence,  $\tilde{f}^R - \tilde{f}^R(x^*) \geq 0$  leads to  $\tilde{f}^R(x^*) \leq \tilde{f}^R; \forall x$ . Then,  $x^*$  is a nearly possibly optimal solution for RP.

TABLE 2: Comparisons of different researchers' contributions.

Author's name	Weighting method	$\varepsilon$ - constraint method	KKT optimality	Efficient solution	Parametric study	Environment
Khalifa [30]	√	×	×	√	×	Rough set
Osman et al. [25]	×	×	×	√	×	Fuzzy set
Ammar and Emsimir [31]	√	×	√	√	×	Fuzzy set
Ahmed [35]	×	×	√	√	×	Fuzzy set
Ammar and Al- Al- Asfar [32]	√	×	√	√	×	Fuzzy set
Our investigation	√	√	√	√	√	Rough set

If  $x^* \in S^U, x^* \notin S^L$ ; then,  $f^{UAI}(x^*) \leq f^{UAI}(x); \forall x$ , and we have

$$\begin{aligned} \tilde{f}^R(x) - \tilde{f}^R(x^*) &\geq \frac{\gamma_{m+2}^*}{(1 - \gamma_{m+1}^* + \gamma_{m+2}^*)} f^{UAI}(x^*) \\ &\quad - \frac{\gamma_{m+1}^*}{(1 - \gamma_{m+1}^* + \gamma_{m+2}^*)} f^{LAI}(x^*). \end{aligned} \quad (45)$$

From KT conditions, we have

$$\tilde{f}^R - \tilde{f}^R(x^*) \geq \frac{\gamma_{m+1}^*}{(1 - \gamma_{m+1}^* + \gamma_{m+2}^*)} f^{LAI}. \quad (46)$$

From the assumption that  $f^{LAI}(x^*) > 0$ , and  $x^*$  is not solution for BP,

$$\frac{\gamma_{m+1}^*}{(1 - \gamma_{m+1}^* + \gamma_{m+2}^*)} = 0. \quad (47)$$

Thus,  $\tilde{f}^R(x) - \tilde{f}^R(x^*) \geq 0$ , which implies to  $\tilde{f}^R(x^*) \leq \tilde{f}^R(x); \forall x$ . Then,  $x^*$  is a nearly possibly optimal solution for RP.  $\square$

## 6. Numerical Example

We consider the following problem  $\tilde{f}^R(x): X \rightarrow \mathbb{R}$  with  $f^{LAI}(x) = x_1 + x_2$ ,  $f^{UAI}(x) = 1/3x_1^3 - 2x_1^2 - 10x_2 + 100$  and consider the following RP problem as

$$(RP) \min \tilde{f}^R(x). \quad (48)$$

Subject to

$$X = \{(x_1, x_2) \in \mathbb{R}^2: x_1 + x_2 \leq 10, 3.5 \leq x_1 \leq 6, x_2 \leq 6, x_1 + x_2 \geq 1\}. \quad (49)$$

Then,

$$(LAP) \min f^{LAI}(x) = x_1 + x_2. \quad (50)$$

Subject to

$$x \in X, \quad (51)$$

$$(UAP) \min f^{UAI}(x) = \frac{1}{3}x_1^3 - 2x_1^2 - 10x_2 + 100.$$

Subject to

$$x \in X. \quad (52)$$

Hence, the BP is

$$(BP) \min F(x) = f^{UAI}(x) - f^{LAI}(x). \quad (53)$$

Subject to

$$x \in X. \quad (54)$$

The solution of the LAP is  $S^L = \{(5, 5)\}$ , and the solution of the UAP is  $S^U = \{(1 - \lambda)(6, 4) + \lambda(4, 6), 0 \leq \lambda \leq 1\}$ . Then,

- (1) There is no one-size-fits-all answer (Definition 9.1)
- (2) The best conceivable solution is (5,5), where  $(5, 5) \in S^L \cap S^U$  and  $F(5, 5) \neq 0$  (Definition 9.2)
- (3) The nearly possibly solution is  $\{(1 - \lambda)(6, 4) + \lambda(4, 6), 0 \leq \lambda \leq 1\} \cup \{(5, 5)\}$  (Definition 9.3)

## 7. Discussion

The proposed approach is compared to some existing literature in this section to show the benefits of the proposed approach. Table 2 investigates this comparison in the case of some parameters

## 8. Concluding Remarks and Future Works

This paper introduces the concept of a rough function, as well as its convexity and differentiability based on its boundary region. The boundary area notion has also been used to investigate a new sort of rough programming challenge and its answers. This research could be expanded to include more fuzzy-like structures in the future (such as interval-valued fuzzy sets, neutrosophic sets, pythagorean fuzzy sets, and spherical fuzzy sets), and more discussion and suggestions could also be included in the future studies. The key features of this study can be summarized as follows:

The proposed study can be extended by developing

- (i) Intuitionistic fuzzy set with a possibility interval value
- (ii) Intuitionistic fuzzy set with a probability interval
- (iii) Fuzzy hypersoft expert set is a possibility
- (iv) Possibility fuzzy pythagorean set
- (v) Possibility picture fuzzy set
- (vi) For example, a spherical fuzzy set

The following are some ideas for further research:

- (1) For rough multiobjective programming, determine the link between rough weights and rough parameters
- (2) An investigation of duality in the context of a rudimentary multiobjective programming problem
- (3) A parametric study of a rough programming issue in which the objective function has roughness
- (4) A parametric investigation of a rough programming problem with rough constraints
- (5) Determine the link between the rough weights and the rough parameter in rough multiobjective programming
- (6) A duality investigation on the problem of rough multiobjective programming
- (7) A parametric analysis of a rough programming problem in which the goal function and restrictions are both rough [36–38]

## Data Availability

No data were used in this manuscript.

## Conflicts of Interest

The authors declare that there are no conflicts of interest.

## Acknowledgments

The researchers would like to express their gratitude to Qassim University's Deanship of Scientific Research for supporting the publishing of this study.

## References

- [1] Z. Ali, T. Mahmood, T. Mahmood, K. Ullah, and Q. Khan, "Einstein geometric aggregation operators using a novel complex interval-valued pythagorean fuzzy setting with application in green supplier chain management," *Reports in Mechanical Engineering*, vol. 2, no. 1, pp. 105–134, 2021.
- [2] Z. Pawlak, K. Slowinski, and R. Slowinski, "Rough classification of patients after highly selective vagotomy for duodenal ulcer," *International Journal of Man-Machine Studies*, vol. 24, pp. 413–433, 1986.
- [3] Z. Pawlak, *Rough Sets*, Kluwer Academic Publishers, Dordrecht, Netherlands, 1991.
- [4] N. Kazemitash, H. Fazlollahabbar, and M. Abbaspour, "Rough best-worst method for supplier selection in biofuel companies based on green criteria," *Operational Research in Engineering Sciences: Theory and Applications*, vol. 4, no. 2, pp. 1–12, 2021.
- [5] T. E. M. Atteya and B. Kamilia, "Convexity in the 1st class of the rough mathematical programming problems," *International Journal of Mathematics Trends and Technology*, vol. 37, no. 3, pp. 224–228, 2016.
- [6] T. Arciszewski and W. P. Ziarko, *Adaptive Expert System for Preliminary Design Wind Bracings in Steel Skeleton Structure, Second Century of the Skyscraper*, Van Nostrand Reinhold Company, New York, NY, USA, 1988.
- [7] H. K. Sharma, K. Kumari, and S. Kar, "Forecasting Sugarcane Yield of India based on rough set combination approach," *Decision Making: Applications in Management and Engineering*, vol. 4, no. 2, pp. 163–177, 2021.
- [8] R. Sahu, S. R. Dash, and S. Das, "Career selection of students using hybridized distance measure based on picture fuzzy set and rough set theory," *Decision Making: Applications in Management and Engineering*, vol. 4, no. 1, pp. 104–126, 2021.
- [9] E. Durmic, Ž. Stević, P. Chatterjee, M. Vasiljević, and M. Tomašević, "Sustainable supplier selection using combined FUCOM – rough SAW model," *Reports in Mechanical Engineering*, vol. 1, no. 1, pp. 34–43, 2020.
- [10] D. Pamucar and A. Jankovic, "The application of the hybrid interval rough weighted Power-Heronian operator in multi-criteria decision making," *Operational Research in Engineering Sciences: Theory and Applications*, vol. 3, no. 2, pp. 54–73, 2020.
- [11] J. Fibak, Z. Pawlak, and K. Slowinski, "Rough sets based decision algorithms for treatment of duodenal ulcer by HSV," *Bulletin of PAS*, vol. 34, no. 34, pp. 227–246, 1986.
- [12] T. Munakata, "Rough control: A perspective," in *Rough Sets and Data Mining Analysis of Imprecise Data*, pp. 77–87, Springer, Berlin, Germany, 1997.
- [13] Z. Pawlak, "Rough sets," in *Rough Sets and Data Mining Analysis of Imprecise Data*, Kluwer Academic, Boston, MA, USA, 1997.
- [14] S. Mitatha, K. Dejhan, F. Cheevasuvit, and W. Kasemsht, "Some experimental results of using rough sets for printed Thai characters recognition," *International Journal of Computational Cognition*, vol. 1, no. 4, pp. 109–121, 2003.
- [15] S. Qiang and C. Alexios, "Combining rough sets and data-driven fuzzy learning for generation of classification rules," *Pattern Recognition*, vol. 32, pp. 2073–2076, 1999.
- [16] W. Jin-Mao, "Rough set based approach to selection of node," *International Journal Computational Cognition*, vol. 1, no. 2, pp. 25–40, 2003.
- [17] R. Slowinski, *Intelligent Decision Support-Handbook of Advances and Applications of the Rough Set Theory*, Kluwer Academic Publishers, Dordrecht, Netherlands, 1992.
- [18] B. Liu, *Rough Programming. Theory and Practice of Uncertain Programming*, Springer, Berlin, Germany, 2002.
- [19] P. Lv and P. Chang, "Rough programming and its application to production planning," in *Proceedings of the International Conference on Risk Management & Engineering Management (IEEE Explore)*, Beijing, China, November, 2008.
- [20] M. S. Osman, E. F. Lashein, E. A. Youness, and T. E. M. Atteya, "Mathematical programming in rough environment," *Optimization*, vol. 60, no. 5, pp. 603–611, 2011.
- [21] R. H. Waliv, U. Mishra, H. Garg, and H. P. Umap, "A Nonlinear programming approach to solve the stochastic multi-objective inventory model using the uncertain information," *Arabian Journal for Science and Engineering*, vol. 45, no. 8, pp. 6963–6973, 2020.
- [22] S. Agarwal, S. S. Dandge, and S. Chakraborty, "Parametric analysis of a grinding process using the rough sets theory," *Facta Universitatis – Series: Mechanical Engineering*, vol. 18, no. 1, pp. 91–106, 2020.
- [23] D. Bozanic, A. Randjelovic, M. Radovanovic, and D. Tesic, "A hybrid LBWA - IR-MAIRCA multi-criteria decision-making model for determination of constructive elements of weapons," *Facta Universitatis Series: Mechanical Engineering*, vol. 18, no. 3, pp. 399–418, 2020.
- [24] J. Xu and L. Yao, "A class of multiobjective linear programming models with random rough coefficients," *Mathematical and Computer Modelling*, vol. 49, pp. 189–206, 2009.

- [25] M. S. Osman, M. El- Sherbiny, H. A. Khalifa, and H. H. Farag, "A fuzzy technique for solving rough interval multiobjective transportation problem," *International Journal of Computer Application*, vol. 147, no. 10, pp. 49–57, 2016.
- [26] T. E. M. Attaya, "Rough multiple objective programming," *European Journal of Operational Research Elsevier*, vol. 248, no. 1, pp. 204–210, 2016.
- [27] M. G. Brikaa, Z. Zheng, and E.-S. Ammar, "Fuzzy multi-objective programming approach for constrained matrix games with payoffs of fuzzy rough numbers," *Symmetry (MDPI- AG)*, vol. 11, no. 5, 2019.
- [28] S. Ghosh and S. K. Roy, "Fuzzy- rough multiobjective product blending fixed- charge transportation problem with truck load constraints through transfer station," *RAIRO - Operations Research*, vol. 55, Article ID S2923, 2021.
- [29] F. Ahmad, S. Ahmad, A. T. Soliman, and M. Abdollahian, "Solving multi- level multiobjective fractional programming problem with rough interval parameters in neutrosophic environment," *PAIRO Operations Research*, vol. 55, pp. 2567–2581, 2021.
- [30] H. A. Khalifa, "Study on multi- objective nonlinear programming in optimization of the rough interval constraints," *International Journal of Industrial Engineering& Production Research*, vol. 29, no. 4, pp. 407–413, 2018.
- [31] E. E. Ammar and A. A. A. Emsimir, "A mathematical model for solving integer linear programming problems," *African Journal of Mathematics and Computer Science Research*, vol. 13, no. 1, pp. 39–50, 2020.
- [32] E. E. Ammar and A. Al- Al- Asfar, "Approaches for solving fully fuzzy rough multi- objective nonlinear programming problems," *Academic Journal of Applied Mathematical Sciences*, vol. 7, no. 2, pp. 113–128, 2021.
- [33] Z. Liao, X. Zhang, Q. Zhang, W. Zhang, and W. Li, "Rough approximation- based approach for designing a personalized tour route under a fuzzy environment," *Information Sciences*, vol. 575, pp. 338–354, 2021.
- [34] F. A. Farahat and M. A. ElSayed, "Achievement stability set for parametric rough linear goal programming problem," *Fuzzy Information and Engineering*, vol. 11, no. 3, pp. 279–294, 2021.
- [35] F. Ahmed, "Robust neutrosophic programming approach for solving intuitionistic fuzzy multi-objective optimization problem," *Complex& Intelligent Systems*, vol. 7, 2021.
- [36] V. Chakong and Y. Y. Haimes, *Multi- Objective Decision Making Theory and Methodology*, Dover Publications, Inc, North- Holland, New York, NY, USA, 1988, <https://lib.ugent.be/catalog/rug01:000485253>.
- [37] H. Garh, S. S. Alodhaibi, and H. A. Khalifa, "Study on multi-objective nonlinear programming problem with rough parameters," *Journal of Intelligent and Fuzzy Systems*, vol. 42, no. 4, pp. 3591–3604.
- [38] S. Midya and S. K. Roy, "Multi- objective fixed charge transportation problem using rough programming," *International Journal of Operational Research*, vol. 37, no. 3, 2020.

## Research Article

# Decision Scheduling for Cloud Computing Tasks Relying on Solving Large Linear Systems of Equations

Jing He 

*College of Artificial Intelligence and Big Data, Chongqing Industry Polytechnic College, Chongqing, China*

Correspondence should be addressed to Jing He; [hejing@cqipc.edu.cn](mailto:hejing@cqipc.edu.cn)

Received 20 January 2022; Accepted 21 February 2022; Published 19 March 2022

Academic Editor: Daqing Gong

Copyright © 2022 Jing He. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the continuous reform and innovation of Internet technology and the continuous development and progress of social economy, Big Data cloud computing technology is more and more widely used in people's work and life. Many parallel algorithms play a very important role in solving large linear equations in various applications. To this end, this article aims to propose and summarize a cloud computing task scheduling model that relies on the solution of large linear equations. The method of this paper is to study the technology of solving large-scale linear equations and propose an M-QoS-OCCSM scheduling model. The function of the experimental method is to solve the problem of efficiently executing  $N$  mutually dependent parallel tasks within limited resources, while fully satisfying users' expectations of task completion time, bandwidth rate, reliability, and cost. In this paper, the application experiment of large-scale linear equations in task scheduling is used to study task scheduling algorithms. The results show that when the task load is 10 and 20, the convergence speed of the MPQGA algorithm is 32 seconds and 95 seconds faster than that of the BGA algorithm, respectively.

## 1. Introduction

The rapid development of the Internet has also promoted the development of cloud computing, and people's research on cloud computing has become more and more in-depth and extensive. Large-scale data mining and distributed processing technology on the Internet have also been paid more and more attention. This paper combines the characteristics of the two and proposes a model combining text mining and task scheduling under cloud computing. Cloud computing is a new business computing model that uses network connections to obtain various applications, data, and IT services. The core of cloud computing is the integrated scheduling and management of cloud environment and task resources submitted by users according to user needs, and users only need to pay on demand. Therefore, for cloud services, how to meet the different needs of different users for quality of service (QoS) is an important issue that needs to be considered when scheduling cloud computing.

The study of resource deployment and task scheduling problems and their solutions in the cloud computing

environment is of great significance and value for the theoretical research and application practice of cloud computing. As a business service, cloud computing must not only consider optimizing task scheduling strategies to improve the service capabilities of the system but also consider the service revenue of cloud service providers. There are some controversies in this area, but there is no efficient solution yet. Therefore, studying cloud computing task scheduling strategies has important theoretical value and practical significance for improving the service capabilities of cloud computing systems.

This paper proposes a balanced clone scheduling algorithm that can effectively improve resource utilization and task scheduling efficiency. This paper proposes a model that combines data mining and task scheduling in a cloud environment. Combining the advantages of data mining and distributed computing, this model can calmly deal with data processing of massive amounts of information and provide efficient services for users. This paper comprehensively considers the transmission cost, processing time, processing cost, and transmission time of data in the cloud computing



environment, and proposes a mathematical model for task scheduling optimization. And this paper proposes a mathematical model and particle swarm optimization algorithm design based on the increase of the variable neighborhood. The simulation results show that the proposed optimization model and optimization algorithm can not only optimize the time, but also optimize the cost.

## 2. Related Work

The road to solving large-scale linear equations has never stopped. In recent years, people have gradually combined it with task scheduling in order to achieve unexpected results. Wang et al. propose an SSLE method for general constrained optimization based on the mixed exact penalty function. He added automatic adjustment rules to the algorithm for the selection of penalty parameters to ensure that the number of updates of penalty parameters is limited. He also extended the Facchinei activity set recognition technology to general constrained optimization and gave the corresponding recognition function. In each iteration, the algorithm only solves two or three simplified linear equations with the same coefficients to obtain the search direction. Without assuming strict complementarity and weaker than strong second-order sufficient conditions, the convergence rate of the algorithm is proved to be super linear [1]. Phuc and Verbitsky obtained the global solution of quasi-linear equations with metric coefficients and solved a large number of model problems [2]. Lv and Wang studied the existence, uniqueness, and asymptotic stability of traveling wave fronts for discrete quasilinear equations with time delays. He first established the existence of traveling wave fronts using hyperon solutions and monotonic iteration techniques. Then he proved that the traveling wave front is unique before the translation. Finally, he used the comparison principle and compression technique to prove that the traveling wave front has a phase shift and is globally asymptotically stable [3]. Mingtong introduced the basic process of automatic modeling of large buildings. By extending the L-system theory, using elevation maps and automatic river recognition, he researched and proposed an automatic road generation method for complex terrain. Experiments show that the results given by his method are consistent with the surrounding terrain environment. It can automatically generate modifiable roads in complex terrain constrained by altitude and rivers. The road generation algorithm, based on the Voronoi diagram he proposed, can effectively make up for the lack of road changes generated by the system [4]. Samriya et al. proposed a multiobjective penguin optimization algorithm. The proposed method is analyzed through binary gravity search algorithm, ant colony optimization, and particle swarm optimization, which makes it suitable for virtual machines in data centers. Compared with other strategies, the algorithm he proposed is energy-efficient and has significant differences [5]. Jabir et al. proposed an enhanced antlion optimization algorithm mixed with the popular particle swarm optimization algorithm to optimize workflow scheduling specifically for the cloud. The research aims to provide enhanced workflow scheduling that is safer

than the existing framework. It enhances the program's ability to evaluate according to cost, load, and completion time [6]. Tang et al. proposed an energy-saving workflow task-scheduling algorithm based on DVFS, whose purpose is to obtain more energy savings and maintain service quality on the premise of meeting deadlines. This algorithm can recycle useful idle time after the server is merged and obtain the entire makespan and deadline based on the heterogeneous-earliest-finish-time algorithm [7]. Gruzlikov et al. proposed a pipeline workshop scheduling method for the calculation process in a distributed real-time system. This method is based on a solvable system with a simple optimization scheduling algorithm. This method opens up ideas for improving the efficiency of the assembly line workshop [8, 9].

## 3. Linear Equations Solution Method and Task Scheduling Method

*3.1. Common Large-Scale Linear Equation Solving Technology.* There are two methods for solving linear equations in direct method and iterative method [10]. Small and medium-sized equations ( $n < 10000$ ) are often solved by the direct method, and the direct method needs to decompose the coefficient matrix  $A$ . Under normal circumstances, matrix sparsity cannot be guaranteed. As the degree  $n$  increases, the amount of memory required becomes  $O(n^2)$ , and the amount of floating-point operations required becomes  $O(n^3)$ . The sparse direct method based on sorting has some control over memory requirements and time consumption, allowing modern single-CPU personal computers to solve 100,000-order matrices. However, the success of the sparse direct method is related to the nature of the matrix itself, and higher-order linear systems are powerless [11]. Therefore, iterative methods are often used for large or very-large equations. In simple terms, the iterative method is to create a modified  $S_k$  for the initial approximate solution vector  $X_0$  of each step. This makes the next approximate solution vector  $X_{k+1}$  equal to the current approximate solution plus the correction vector  $S_k$ .

$$X_{k+1} = S_k + X_k. \quad (1)$$

This produces an iterative vector sequence  $\{X_k\}$ . It can be seen that the ideal correction vector should be

$$S_k = X^* - X_k = Q(b - Q \cdot r_k), \quad (2)$$

where  $X^*$  is the solution of the simultaneous equations, and  $r_k$  represents the  $k$ th residual vector. However, in order to obtain this ideal correction vector, this article needs to solve equation  $Q \cdot r_k = X_k$ , which is actually the same problem as solving the original equation. Different iterative methods represent different methods of calculating the correction vector  $S_k$  [12]. The classic iterative method is based on matrix division, assuming that the coefficient matrix  $A$  has the following divisions.

$$A = P - Q, \quad (3)$$

where  $P$  is the invertible matrix and the original system of equations is  $(P - Q)x = b$ . Therefore, the following iteration sequence is constructed.

$$X = P^{-1}Q_{xk} + P^{-1}b. \quad (4)$$

In order to avoid calculating  $P^{-1}Q_{xk}$  and  $P^{-1}b$ , it can turn to solving the equations in the following way:

$$P_{xk} = Q_{xk} + b. \quad (5)$$

Different  $P$  choices have different segmentation methods and different iterative methods [13]. Assumptions:

$$A = D - C_L - C_K. \quad (6)$$

Among them,  $D$  is the diagonal part of  $A$ ,  $C_L$  is the lower triangular part of  $A$ , and  $C_K$  is the upper triangular part of  $A$ .

In order to achieve better results, people constructed a symmetric over-relaxation method (SSOR method), splitting  $A$  into  $A = P_{ssos} - Q_{ssor}$ , where

$$P_{SSOR} = \frac{w}{2-w} \left( \frac{1}{w} D - C_L \right) D^{-1} \left( \frac{1}{w} D - C_K \right). \quad (7)$$

In addition, there are methods such as alternating direction iteration, Hermitian and anti-Hermitian splitting, and multiple splitting, all of which are based on practical problems and are all numerically algebraic [14]. The coefficient matrix of the linear equation system obtained by the finite-element method or the method of moments is generally a complex symmetric matrix, which can be described in the following format.

$$(A + iB)(z + iy) = b + ic. \quad (8)$$

Among them,  $A, B \in R$  is a real symmetric matrix. The Krylov subspace method of common asymmetric matrices (such as GMRES, BICGSTAB, and QMR) can be used to solve complex symmetric linear equations. However, using these methods to solve the problem does not take advantage of the symmetry of the matrix. Some solutions transform the symmetric system into an equivalent real symmetric system and then use the Krylov method to solve the real symmetric matrix [15]. The following are some iterative methods for directly solving (8). Using the concept of quasi inner product  $(x, y) = x^T y$ , where the  $x, y \in C$  superscript letter  $T$  represents transpose instead of conjugate transpose, the QMR method can recursively reach the minimum in a short time. This type of quantity and special matrix are often encountered in actual engineering [16]. The key to the success of the Krylov subspace method is the choice of prerequisites. The preprocessing is to transform a linear equation system into another linear equation system with the same solution, but the transformed equation system has the characteristic of facilitating iterative solution. The prerequisite is to generate a matrix for this transformation. If  $P$  is an invertible matrix somewhat similar to  $A$ , then the preprocessed system looks like the following equation:

$$P^{-1}A_x = P^{-1}b. \quad (9)$$

If there is the same solution as equation (1), and there is an equation that is easy to solve, then  $P$  is called a prerequisite. Equation (9) is preprocessed from the left side and can also be preprocessed from the right side.

$$AP^{-1}y = b, \quad x = P^{-1}y. \quad (10)$$

Or, preprocess both sides at the same time:

$$P_1^{-1}AP_2^{-1}y = P_1^{-1}b, \quad x = P_2^{-1}y. \quad (11)$$

Then, the precondition in formula (11) is  $M = M_1M_2$ . A good precondition should have the following two properties: the system after the precondition should be easy to solve, the precondition should be easier to construct and the cost of applying the precondition should not be too high [17].

In modern computer architecture systems, the biggest problem with incomplete decomposition of prerequisites is that it is not easy to parallelize the process of construction and use. Approximate prerequisites for adapting to modern computer structure systems have appeared, and their implementation has natural parallelism [18]. In recent years, the algebraic multigrid (AMG) method that combines the area division (DD) method, multigrid (MG), and physical background to solve partial differential equations (PDEs) has become a hot spot in numerical computing [19].

**3.2. M-QoS-OCCSM Scheduling Model Based on Cloud Computing.** The architecture of the M-QoS-OCCSM collaborative scheduling model proposed in this chapter is shown in Figure 1.

As shown in Figure 1, the execution process of the entire M-QoS-OCCSM scheduling model is as follows. First, the user submits the application task with the deadline of the bottom line and the scheduling budget target constraint to the task scheduler. Then the task scheduler sends the QoS target constraint conditions and other related parameter information of the application task to the extension module of the M-QoS-OCCSM model [20]. Next, the extension module of the M-QoS-OCCSM model first applies the method described in the previous section to model the application task with multiobjective constraints. Then comprehensively consider the application task's QoS target constraints and user expectations and based on a membership function, apply the method described to convert the application task's multiple QoS target constraints into a single-objective constraint optimization problem. Finally, the reconstructed genetic algorithm is applied to approximate the optimal solution of the mentioned single-objective optimization problem [21]. The calculated result is the final scheduling decision plan, which is returned to the scheduler. The scheduler selects appropriate computing resources for the application tasks for scheduling according to the optimization results, thereby completing the scheduling decision.

The extension module serves as an integrated unit in the entire M-QoS-OCCSM scheduling model architecture. For some applications, it can be designed as a separate middleware module for real task-scheduling scenarios [22].

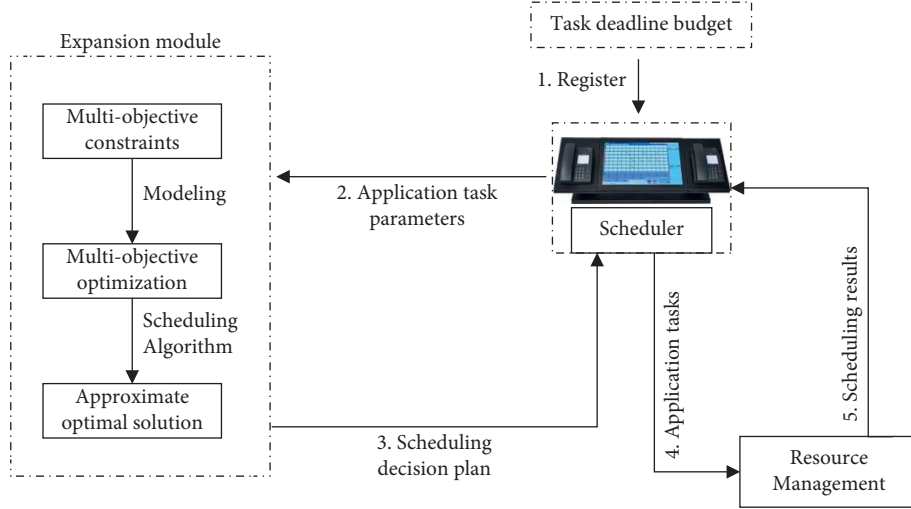


FIGURE 1: M-QoS-OCCSM system architecture.

From the user's point of view, the user always hopes to complete all application task scheduling requests within a desired time interval. This means that the actual execution of application tasks should be reduced as much as possible on the assigned computing nodes. At the same time, for a specific resource  $R$  available on a computing node, the total computing power consumed by all application tasks on the resource  $R$  cannot exceed the computing power that the resource can provide [23].

It can be seen from the analysis that the QoS target constraint of the aforementioned time limit bottom line can be expressed as follows:

$$\begin{aligned} \min(O_{\text{deadline}}) &= \sum_{t=1}^T \sum_{n=1}^N M_{tn} V_{tn}, \\ \text{s.t. } \sum_{t=1}^T \sum_{n=1}^N \sum_{k=1}^K M_{tn} R &\leq \sum_{n=1}^N \sum_{k=1}^K S(CN_{tn}). \end{aligned} \quad (12)$$

On the contrary, users always want to save scheduling costs as much as possible while obtaining satisfactory services. Therefore, the cloud system should complete the scheduling and execution of all tasks as much as possible under the constraints of the established scheduling budget [24]. In other words, in an established cloud computing environment, the system should reduce the total scheduling cost of all tasks as much as possible. In addition, for a specific application task  $t$ , the sum of the scheduling costs consumed by all available resources on the target computing node cannot exceed the task scheduling budget.

According to the analysis, the QoS target constraint of the application task scheduling budget can be expressed as follows:

$$\min(O_{\text{budget}}) = \sum_{t=1}^T \sum_{n=1}^N M_{tn} C_{tn}, \text{ s.t. } \sum_{n=1}^N M_{tn} C_{tn} \leq B_n, \forall n \in T. \quad (13)$$

This paper compares the algorithm based on QoS constraints with the basic genetic algorithm. The relationship between tasks with interdependence constraints is shown in Figure 2, and the specific parameters and weight vector settings are shown in Table 1.

The user wants to pay as little scheduling cost as possible for the scheduling request of the application task. In real-world cloud computing application scenarios, the cost of scheduling application tasks is often directly proportional to the user's satisfaction with completing the scheduling goals. Therefore, users want to spend less on their application task scheduling budget, assuming they have a satisfactory scheduling service [25]. As shown in formula (14), assuming that a satisfactory scheduling service is obtained, the membership function of the task scheduling expenditure can be used to express the membership of the task scheduling budget and the scheduling target.

$$Y_{nm}^{\text{budget}} = 1 - \frac{C_{\min}}{C_{nm}}. \quad (14)$$

In the formula,  $C_{\min}$  represents the minimum expected value of the scheduling overhead of the application task  $t_n$ , and  $C_{nm}$  represents the scheduling budget when the application task  $t_n$  is scheduled on the target computing node  $C_{nm}$ . It can be seen from equation (14) that as the scheduling budget expenditure  $C_{nm}$  of the application task  $t_n$  decreases, the membership function value  $Y_{nm}^{\text{budget}}$  of the scheduling budget decreases. Since users want to pay as little task scheduling budget as possible, the goal of the final task scheduling solution is to minimize the value of the mentioned membership function as much as possible [26].

Through the analysis of the two QoS target constraints, deadline and scheduling budget. This article finds that with less scheduling budget and the shortest possible deadline, everything needed to complete an application task scheduling request has increased. It tries to minimize the aforementioned period. The membership function value of time and the membership function value of scheduling budget expenditure. Based on this conclusion, two membership

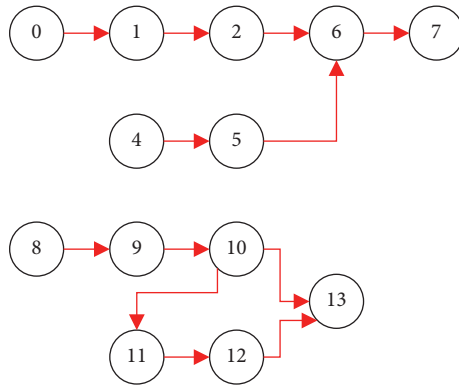


FIGURE 2: Interdependence diagram between tasks.

TABLE 1: Task scheduling algorithm parameter setting table.

	Two task scheduling algorithms
Total resources	15
Total number of tasks	45
Unit price of virtual machine resources	7
Virtual machine computing performance (million instructions per second)	1500
Task length (one hundred instructions)	10,000
Virtual machine resource bandwidth (Mb)	1024
Expected time (seconds)	65
Expected bandwidth (Mb)	910
Expected cost	4500
Number of generations	125
Crossover probability parameter	$k1 = 0.6, k2 = 0.8$
Mutation probability parameter	$k3 = 0.2, k4 = 0.06$
Variation gene range parameter, $d$	5

functions can be used to transform the task scheduling multi-QoS objective constraint optimization problem into a single-objective constraint optimization problem. The transformed single-objective constrained optimization problem is shown as follows:

$$\min(O) = w_1 \min(O_{\text{deadline}}) + w_2 \min(O_{\text{budget}}). \quad (15)$$

In the formula, the deadline and scheduling budget weight parameters  $w_1$  and  $w_2$  satisfy  $\sum_{n=1}^2 w_n = 1, 0 \leq w_1 \leq 1$ . When scheduling tasks, pay attention to the deadlines and the scheduling budget constraints of each user. Therefore, when constructing the objectives of the mentioned task scheduling scheme, adjusting the two weight parameters for different users can meet the various QoS target constraints of users.

By default, Hadoop usually adopts a first-come first-serve strategy. The advantages of this strategy are simplicity and low overhead, while reducing the burden of the job tracker [27]. The basic idea of solving the problem is to queue up all tasks submitted by the end customer according to the time when the job was submitted. The execution order of the job queue is usually determined by priority and transmission order. That is, the job submitted first by the system has a higher priority by default and will be processed earlier. But the disadvantage of the FIFO strategy is that it is

unfair. According to the solution idea of this strategy, for those jobs with low priority, the chance of being processed will be greatly reduced, and the waiting time for idle machines will be particularly long. That is to say, it is difficult to guarantee the QoS for jobs with low priority. This strategy lacks consideration of differences in job requirements.

**3.3. Cloud Computing Task Scheduling Strategy.** Cloud computing resources include storage resources, computing resources, network resources, and so on. In fact, these resources are abstracted into services through virtualization technology and provided to the outside world. On the contrary, there is a correlation between service quality and resource occupancy rate, service quality, and energy consumption. Therefore, when optimizing resources, this article must consider comprehensively, rather than biasing one aspect. Therefore, when optimizing resource deployment and scheduling, all resources must be fully considered and jointly managed and optimized. It is necessary to optimize the QoS, while optimizing energy consumption and cost [28].

The state of task scheduling in the cloud computing environment can be explained as follows. The total number of resources is  $P$ , and the corresponding set is  $R = \{r1, r2 \dots rn\}$ . The total number of jobs submitted by



users is  $M$ , corresponding to set  $J = \{J1, J2 \dots JM\}$ . Assuming that the  $M$  jobs corresponding to the set  $T = \{T1, T2 \dots TN\}$  are divided into  $N$  tasks, the  $Jm$ th job is divided into  $T_{Num}(Jm)$  tasks, and then the total number of tasks corresponding to the total number of tasks is as follows:

$$N = T_{total}(M) = \sum_{m=1}^M T_{Num}(Jm). \quad (16)$$

Network bandwidth is an important indicator to measure network usage [29]. The size of the bandwidth determines the size of the network transmission capacity, which in turn affects the communication efficiency in the cloud environment. The more frequent the communication, the greater the amount of information, and the higher the bandwidth requirements. Let  $B_{wm}$  be the resource bandwidth of the cloud computing environment,  $B_{user}$  represents the expected bandwidth of the job  $Jm$  specified by the user, and  $B_i$  represents the expected bandwidth of the task  $T_i$  divided by the job  $Jm$ , then

$$B_{user} = \sum_{m=1}^{T_{Num}(Jm)} B_i. \quad (17)$$

The function of user satisfaction  $w$  obtained bandwidth is as follows:

$$W = \left[ \frac{\theta}{T_{Num}(Jm)} \right] \sum_{m=T_{total}(M-1)}^{T_{total}(M)} \ln \left( \frac{B_{wm}}{B_i} \right). \quad (18)$$

Assuming that the resource failure rate in the cloud computing environment is  $P$  (which can be obtained through the resource monitoring system) and the user's expected task completion rate is  $P_{succ}$ , the user satisfaction function of the job completion rate is as follows:

$$W_{succ}(Jm) = \theta \ln \left[ \frac{(1-p)}{P_{succ}} \right]. \quad (19)$$

Cost constraint is one of the most popular QoS constraints, and cost is one of the important components of user QoS. Assuming that resources are billed per unit,  $P_i$  is the number of resources, and  $C_{cpu}$ ,  $C_{men}$ ,  $C_{stor}$ , and  $C_{BW}$  represent the prices of CPU, memory, storage, and bandwidth resources. Then, the total cost of task  $T_i$  can be expressed as follows:

$$C_i = P_1 C_{cpu} + P_2 C_{men} + P_3 C_{stor} + P_4 C_{BW}. \quad (20)$$

Through the previous related research on the existing results of the equations solution, it can be found that each equation solution must have a suitable data center network structure, such as a centralized structure, a distributed structure, and a hierarchical structure [30]. In view of this, before introducing the DRMS equation solution introduced by the task scheduling system architecture in the cloud computing environment, first introduce the data center network structure that implements the equation solution.

As shown in Figure 3, the data center that implements the DRMS equations solution is a hierarchical structure. This

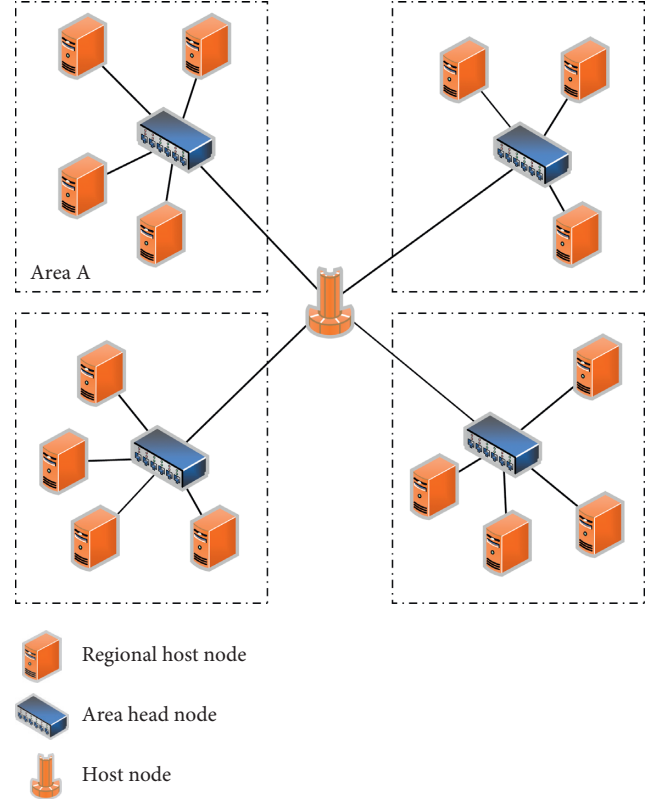


FIGURE 3: Cloud computing data center network structure diagram.

structure is mainly composed of three types of nodes: regional host node, regional head node, and ordinary host node (also called site, or Site) [31].

## 4. Task Scheduling Experiment and Analysis Based on Cloud Computing

**4.1. Application of Large-Scale Linear Equations in Task Scheduling.** In mathematics, the Gaussian elimination method is also known as the Gauss Jordan elimination method, which is an algorithm for solving linear algebra system problems using linear equations [32]. The algorithm first determines the rank of the matrix, and then calculates the inverse of the invertible square matrix. When designing the Gauss method that can cancel the inverse allocation process, the elimination method also reduces the unknown coefficients on the main diagonal to zero. The elimination method does not need to replace the diagonal coefficient matrix equation, this method is called Jordan elimination method. However, the number of multiplication and division operations in Jordan elimination is twice that of pure Gaussian elimination. This is not desirable when solving linear equations, but it is useful when solving inverse matrices.

The operating speeds of Gauss elimination method, and Jordan elimination method are shown in Tables 2 and 3, respectively.

The experimental results are shown in Table 4 by running the Gauss elimination method on a 4-core

TABLE 2: Experimental data of Gauss elimination method.

Number of parallel nodes	Dense matrix scale	1002	2002	5002	6002	9002
1	Total time	0.053	0.0817	0.0093	0.1104	0.0523
	Communication time	0.0576	0.0994	0.0935	0.099	0.0254
	Parallel computing time	0.0446	0.0064	0.0344	0.1106	0.0059
	Speedup ratio	—	—	—	—	—
	Parallel efficiency	—	—	—	—	—
2	Total time	0.1128	0.0884	0.0534	0.0554	0.0488
	Communication time	0.0777	0.0155	0.0955	0.1121	0.0932
	Parallel computing time	0.0574	0.0069	0.0928	0.1062	0.0052
	Speedup ratio	0.67	0.53	0.31	0.85	0.85
	Parallel efficiency	0.33	0.92	0.99	0.39	0.49
4	Total time	0.1069	0.0932	0.0617	0.0501	0.0181
	Communication time	0.0988	0.0045	0.0873	0.0056	0.044
	Parallel computing time	0.0101	0.0868	0.0439	0.0328	0.0838
	Speedup ratio	1.12	0.9	0.88	0.98	0.68
	Parallel efficiency	0.4	1.19	0.51	0.3	0.64

TABLE 3: Experimental data of Jordan elimination method.

Number of parallel nodes	Dense matrix scale	2002	4002	5002	6002	10002
0	Total time	0.0443	0.1012	0.0583	0.0034	0.1169
	Communication time	0.0676	0.0461	0.1171	0.0917	0.0782
	Parallel computing time	0.006	0.1163	0.1053	0.0995	0.04
	Speedup ratio	—	—	—	—	—
	Parallel efficiency	—	—	—	—	—
3	Total time	0.0699	0.0479	0.0687	0.0279	0.0364
	Communication time	0.048	0.0936	0.0636	0.0055	0.0449
	Parallel computing time	0.099	0.0802	0.0514	0.1047	0.1189
	Speedup ratio	0.39	0.43	0.57	0.45	0.4
	Parallel efficiency	0.55	0.77	0.38	0.61	0.45
4	Total time	0.0935	0.0582	0.1137	0.0265	0.063
	Communication time	0.0772	0.0392	0.0976	0.0854	0.0878
	Parallel computing time	0.119	0.0998	0.033	0.1005	0.0662
	Speedup ratio	0.48	0.62	0.97	0.83	0.7
	Parallel efficiency	0.43	0.63	0.49	0.68	0.73

The comparison of parallel efficiency in various situations is shown in Figure 4.

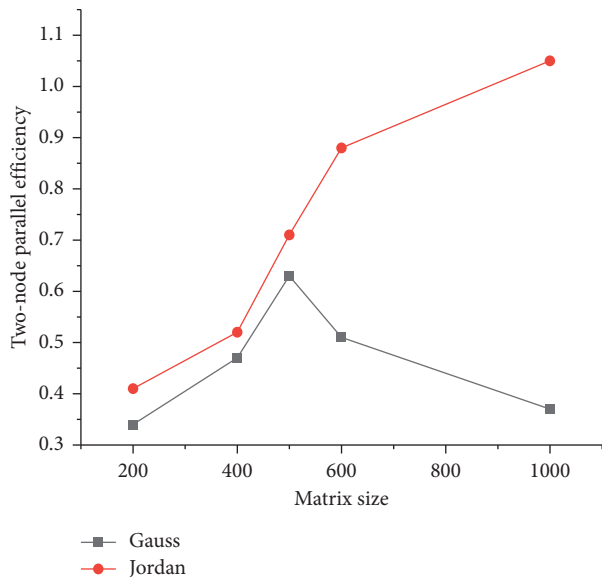


FIGURE 4: Speedup analysis of two elimination methods (two nodes).

multiprocessor to process the Poisson equation. From the data in the table, the speedup for solving the Poisson equation under four cores can be calculated as shown in Figure 5.

The following conclusions can be drawn from analyzing the graphs in this article: ① When the scale of meshing reaches a certain level, compared with the serial solution of single-core, the parallel solution of four-core will highlight its advantage in solution time. Because, when the scale of grid division is small, the communication time between each process takes a larger proportion than the overall running time. Therefore, the parallelism of multiple cores does not show its advantages in the process of solving a smaller meshing scale. ② Analyzing the speedup ratio of various improved parallel algorithms in Figure 5, it can be seen that the nonblocking method can indeed improve the efficiency of parallel solving. And on this basis, overlapping a part of the calculation with nonblocking message passing can further improve the efficiency of parallel solving. ③ From the analysis of the speedup ratio of the parallel algorithm in Figure 5, it can be seen that if the number of processes is



TABLE 4: Poisson equation solving experimental results (unit: second).

Type of equation	$30 \times 60$	$60 \times 120$	$90 \times 180$
Single core nonimproved	0.228	2.396	11.643
Quad-core nonimproved	0.510	2.240	11.648
Quad-core improvement (nonblocking)	0.479	2.442	11.133
Quad-core improvement (overlapping communication)	0.519	2.284	9.620
Gauss elimination method selection times	0.325	2.595	11.882

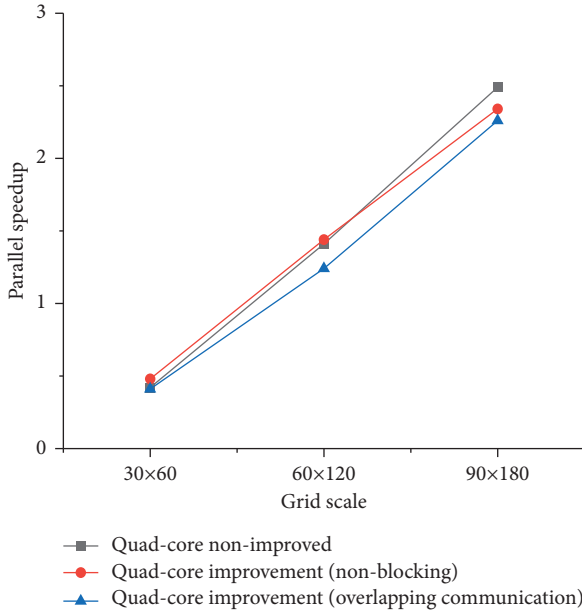


FIGURE 5: Speedup ratio of Poisson equation solution under quad-core.

fixed, the parallel efficiency will increase as the size of the subdomain problem increases. As the problem grows, the number of Gauss-Jordan iterations also increases. Moreover, the increase in the number of iterations of the Gauss elimination method has a linear trend with the increase of the grid division scale, which shows the rationality of solving large-scale sparse linear equations by the Gauss elimination method. This paper also proposes a task scheduling optimization strategy based on service quality optimization, taking into account system load conditions. This article also gives the approximate service cost of the system and its related proofs. Through experiments, this article analyzes the model and TSSQO strategy in detail and compares them with representative strategies. It verifies that the strategy in this article can more reasonably schedule tasks, so that the entire system can not only achieve high service quality, but also relatively low service costs.

**4.2. Multimode Automata Matching Experiment.** Most application services in cloud computing must be decomposed into several subtasks for scheduling, and the decomposed subtasks have different degrees of dependence. How to improve the parallelism, real-time and dynamics of coupling-dependent task scheduling, improve system utilization, and perform reasonable scheduling, and deployment of

dependent task requests has become a hot spot to be solved in current distributed computing and cloud computing.

This experiment tests the time performance of the multimode automata matching method. The tested hardware environment is 200G hard disk, 16G memory, and two 4-core CPUs. The computer operating system is Ubuntu Linux. This experiment tested the matching algorithm on a 1G data set file and compared the time consumption of the traditional algorithm and the multimode automata algorithm when the comparison rules are the same.

It compares the scheduling length and communication overhead performance of TDDPS with HEFT and Min\_Min algorithms in a distributed system composed of three heterogeneous processor nodes with different task sets. And its statistics use TDDPS algorithm to carry on the system load balance that depends on the task scheduling.

The task graph is randomly generated in the experiment. The number of nodes in each task graph [1~100], task dependencies, the amount of communication data between nodes [1~100], and the execution time of each task on different processors [1~100] are all randomly generated. The experiment compares the three scheduling ideas from the three aspects of scheduling length, communication energy consumption, and node resource utilization. Figure 6 shows the scheduling length and communication overhead of the three on different task sets.

According to the results of Figure 6, it can be seen that the TDDPS algorithm has good performance in terms of scheduling length and communication overhead. Owing to the distributed negotiation and scheduling mechanism, each execution node in the heterogeneous system first evaluates the load of the node, and it makes a negotiation response message to the task request according to the real-time load situation of the node. This ensures a certain degree of load balance and improves resource utilization.

In order to test the performance of the BCSOA algorithm, this paper uses the CloudSim simulation platform to simulate and build a data center. At the same time, the algorithm is compared and tested with the DFGA algorithm and the ASAP algorithm. At present, the ASAP algorithm and the DFGA algorithm are typical algorithms in the field of cloud computing task scheduling research. It has achieved relatively ideal experimental results in the specific application process.

The experimental test contains two aspects: the first type of experiment is mainly compared from the perspective of task completion time, and the second type of experiment is mainly compared from the perspective of the balance factor. Figure 7 shows the comparison of the task completion time

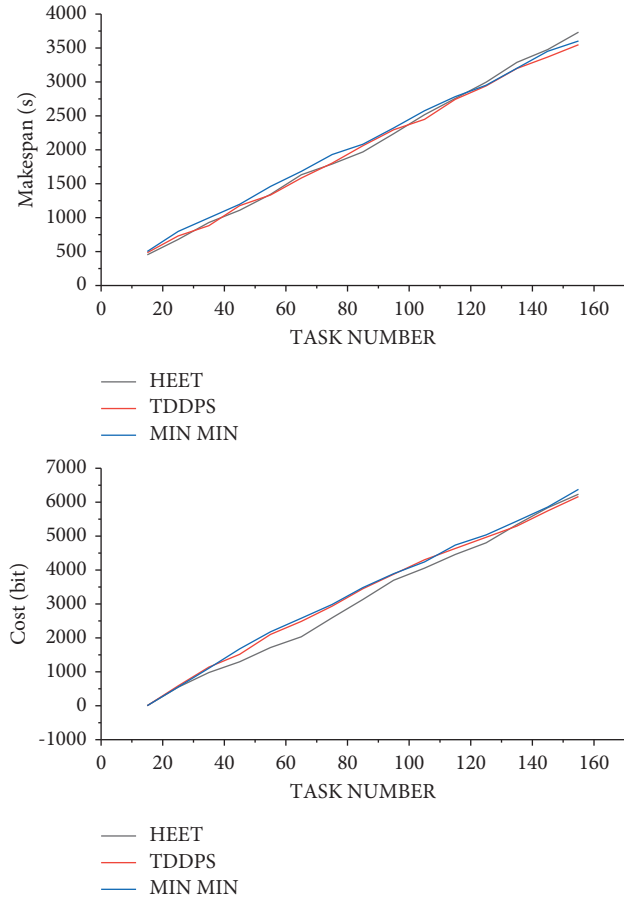


FIGURE 6: Comparison of scheduling length and communication energy consumption of different task sets on three nodes.

of the three task scheduling algorithms when the number of tasks in the system is  $n = 1000$ .

Experimental data show that the traditional matching algorithm has severe time performance degradation when the number of rules increases, and the matching mechanism using multimode automata can greatly improve the matching speed when the rule set is large. Through the comparison of the results of the four rule sets, the time performance of the matching based on the multimodal automata has been greatly improved compared with the traditional algorithm.

**4.3. Task-Scheduling Algorithm.** In order to verify the performance of the MPQGA task-scheduling algorithm, this paper compares the algorithm with the previously mentioned heuristic algorithms HEFT-T, HEFT-B, CPOP, and the basic random search algorithm BGA. This article carried out related tests on two test sets of actual application program graph and randomly generated DAG application table.

The MPQGA algorithm is implemented by C programming. The DAG task graph is represented by a class, and its members include a series of tasks. It uses a two-dimensional matrix to represent the computing speed of each task in a heterogeneous processor and uses another two-dimensional matrix to represent the amount of

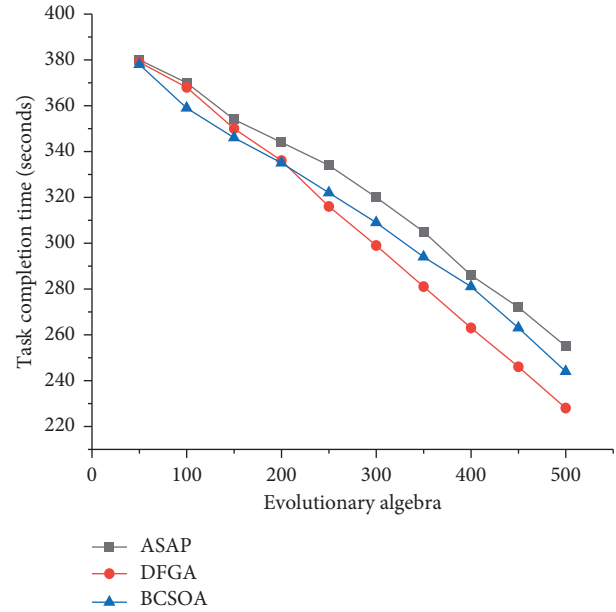


FIGURE 7: Comparison of task completion time of three task scheduling algorithms ( $n = 1000$ ).

communication data between each pair of tasks. The task in the program is also a class, and its members include the predecessor and successor of the task, the input and output of the task, and the amount of data calculation of the task. To implement an FFT with an input vector of  $k$ , there are  $2k - 1$  recursive calls and  $k \log_2 k$  butterfly operations (this article assumes that  $k = 2^i$ ,  $i$  is an integer). In the FFT task graph, each path from the start of the task to the end of the task is a critical path. As the computational cost of tasks at any level is the same, the communication costs of all paths between the two levels are the same. Figure 8(a) shows the average SLR of the task scheduling algorithm for FFT task graphs of different sizes, in which the MPQGA algorithm is better than the average SLR of other algorithms. Figure 8(b) shows the efficiency of different algorithms with 64 data nodes, and the MPQGA algorithm is also better than other algorithms in efficiency.

In the experiment of this article, the value of CCR is set to 0.2, 0.4, 0.6, 1.0, 5.0, and 8.0. Figure 9 shows the average SLR and the efficiency value of the algorithm with different CCR values and different numbers of heterogeneous processors.

Figure 9(a) shows that when the CCR value increases, the average SLR of the algorithm also increases. Figure 9(b) shows the efficiency of the algorithm in the case of different numbers of processors.

In these experiments, this paper uses randomly generated task graphs to evaluate the performance of the algorithm. In order to generate random graphs, this article implements a random graph generator that allows users to generate random graphs with different characteristics. The input parameters of the generator: the number of tasks in the graph, the number of instructions for each task (calculation), the number of subsequent tasks (parallelism), and the CCR value. This article evaluates the performance of the algorithm

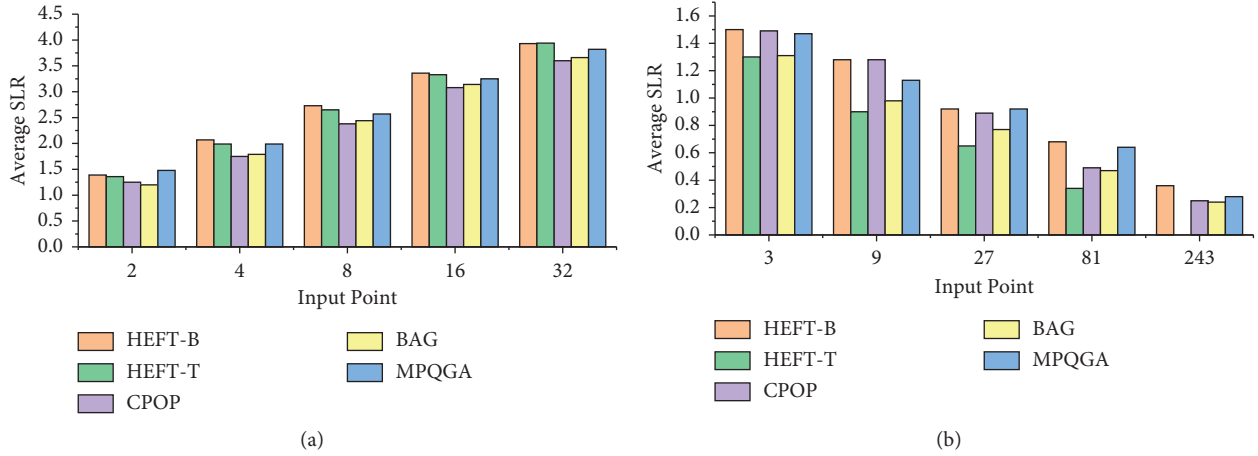


FIGURE 8: Algorithm average SLR and efficiency of fast Fourier transform: (a) algorithm average SLR vs. FFT graph size and (b) algorithm efficiency vs. number of processors.

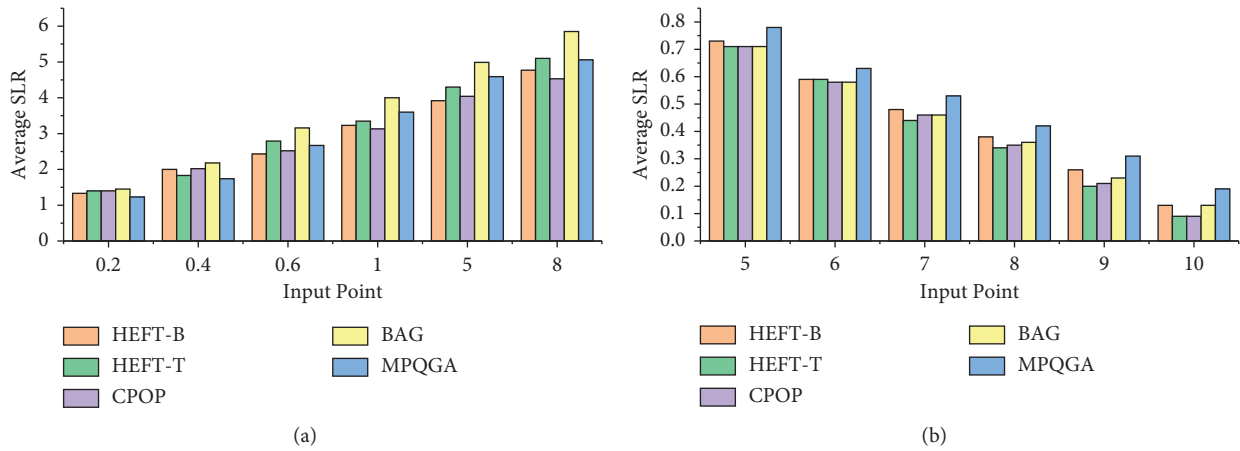


FIGURE 9: Algorithm average SLR and efficiency of molecular dynamics code: (a) algorithm average SLR vs. CCR and (b) algorithm efficiency vs. number of processors.

under different parameters, including different numbers of tasks, different numbers of heterogeneous processors, and different CCR values. The value in Figure 10 is the average maximum and minimum completion time of more than 100 different random DAG graphs.

Figure 10 shows the convergence process of the maximum completion time for a set of randomly generated DAG application graphs (size 10 and 20, respectively). This article can observe that when the DAG application graph is small, the convergence speed of the MPQGA algorithm and the BGA algorithm is very fast. When the DAG application becomes very large, the convergence speed of the algorithm becomes slow, but the convergence speed of the two algorithms is different. The MPQGA algorithm proposed in this paper converges faster than the BGA algorithm, and the final minimum and maximum completion time obtained is better than the BGA algorithm.

In general, with the development of network reliability, cheap computer hardware and software, virtualized hardware technology, and service-oriented architecture, more

and more companies have begun to invest in the development and application of cloud computing technology. More and more companies and individuals are beginning to use cloud computing technology. It continues to promote the progress and development of cloud computing technology. Today, cloud computing, as the core of the next generation of IT technology, has great advantages such as high flexibility, scalability, ease of use, economies of scale, green energy saving, and environmental protection. It is considered to be the next-generation network after the Internet.

## 5. Discussion

Cloud computing has become a trend in the field of IT technology. This article has conducted in-depth research on the task scheduling strategy of cloud computing and has achieved some results. Nevertheless, the work of this article is not yet full and mature enough. At the same time, due to the rapid development of cloud computing technology, there are still many problems in the field of cloud computing

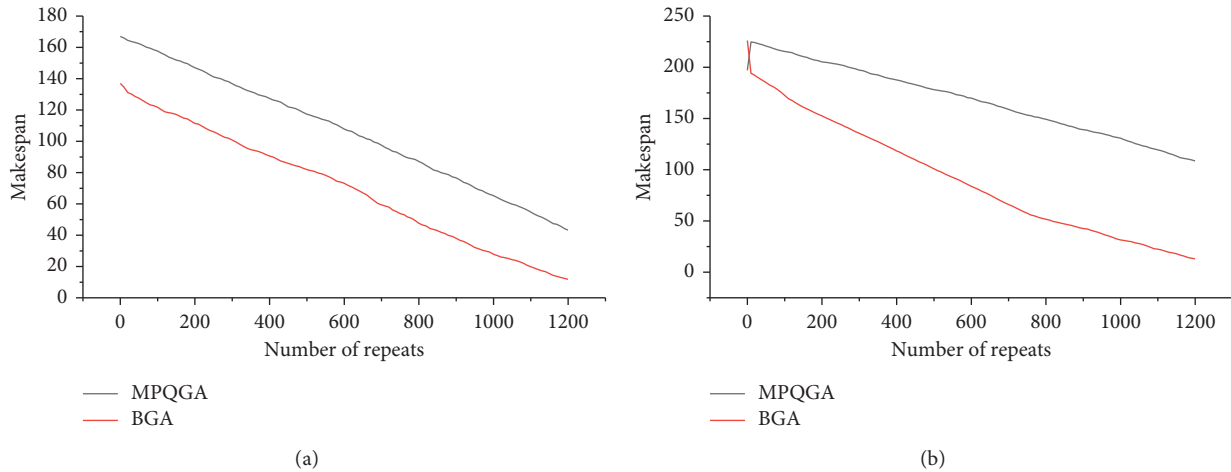


FIGURE 10: The convergence speed of the algorithm under different number of tasks: (a) number of random tasks = 10 and (b) number of random tasks = 20.

technology that need to be further improved and resolved. These problems are mainly reflected in the following aspects: user privacy and data security. The development of cloud computing requires the participation of a large number of users. How to ensure that the user information stored in the cloud is sufficiently safe to ensure that user privacy and user data will not be collected or leaked by cloud service providers and will not be stolen or illegally used by third parties. This is not only a technical issue but also a legal issue. It needs the joint promotion of researchers and government departments. Scholars in the future should strengthen the study of compressed storage of sparse matrices. In this article, the matrix storage of the MPI experiment program is a file that stores all zero elements in the sparse matrix, which will greatly waste the system's memory resources and affect the performance of the parallel program.

## 6. Conclusion

This paper has conducted in-depth analysis and research on the task-dependent scheduling mechanism in cloud computing, including the characteristics and shortcomings of the existing scheduling model, and the problems to be solved by task-dependent scheduling. This paper proposes a new system model suitable for cloud computing environment and a dynamic parallel dependent task scheduling mechanism. In this paper, aiming at the resource-matching problem in the scheduling process, a parallel matching method based on multimode automata is proposed. When the task amount is 10, the calculation speeds of the BGA algorithm and the MPQGA algorithm are 43 and 11, respectively, and when the task amount is 20, the calculation speeds of the BGA algorithm and the MPQGA algorithm are 108 and 13, respectively. It can be seen that the MPQGA algorithm is more excellent, and the more tasks, the more obvious the advantage.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare no conflicts of interest.

## Acknowledgments

This research study was sponsored by Chongqing Yubei District Science and Technology Planning Project. The name of the project is Big Data development of service PAAS automatic delivery and management platform. The authors thank the project for supporting this article.

## References


- [1] Y. Wang, L. Chen, and G. He, "Sequential systems of linear equations method for general constrained optimization without strict complementarity," *Journal of Computational & Applied Mathematics*, vol. 182, no. 2, pp. 447–471, 2017.
- [2] N. C. Phuc and I. E. Verbitsky, "Singular quasilinear and Hessian equations and inequalities," *Journal of Functional Analysis*, vol. 256, no. 6, pp. 1875–1906, 2017.
- [3] G. Lv and M. Wang, "Existence, uniqueness and stability of traveling wave fronts of discrete quasi-linear equations with delay," *Discrete and Continuous Dynamical Systems - Series B (DCDS-B)*, vol. 13, no. 2, pp. 415–433, 2017.
- [4] Li Mingtong, "Research on the mechanism and influence factors of urban style building based on cloud computing logistics information," *Cluster Computing*, vol. 22, no. 6, pp. 13873–13880, 2019.
- [5] J. K. Samriya, S. Chandra Patel, M. Khurana, P. K. Tiwari, and O. Cheikhrouhou, "Intelligent SLA-aware VM allocation and energy minimization approach with EPO algorithm for cloud computing environment," *Mathematical Problems in Engineering*, vol. 2021, no. 6, pp. 1–13, 2021.
- [6] K. Jabir, D. P. Suseelan, and P. M. Keerikkattil, "An effective meta-heuristic based multi-objective hybrid optimization method for workflow scheduling in cloud computing environment," *Cluster Computing*, vol. 24, no. 3, pp. 2367–2384, 2021.
- [7] Z. Tang, L. Qi, Z. Cheng, K. Li, S. U. Khan, and K. Li, "An energy-efficient task scheduling algorithm in DVFS-enabled

- cloud environment,” *Journal of Grid Computing*, vol. 14, no. 1, pp. 55–74, 2016.
- [8] A. M. Gruzlikov, N. V. Kolesov, Y. M. Skorodumov, and M. V. Tolmacheva, “Task scheduling in distributed real-time systems,” *Journal of Computer and Systems Sciences International*, vol. 56, no. 2, pp. 236–244, 2017.
- [9] G. Marković, N. Zdravković, M. Karakašić, and M. Kolarević, “Modified PROMETHEE approach for solving multi-criteria location problems with complex criteria functions,” *Tehnicki vjesnik-Technical Gazette*, vol. 27, no. 1, pp. 12–19, 2020.
- [10] V. Arvind, J. Köbler, S. Kuhnert, and J. Torán, “Solving linear equations parameterized by hamming weight,” *Algorithmica*, vol. 75, no. 2, pp. 322–338, 2016.
- [11] N. V. Dung and V. T. L. Hang, “The generalized hyperstability of general linear equations in quasi-Banach spaces,” *Journal of Mathematical Analysis and Applications*, vol. 462, no. 1, pp. 131–147, 2018.
- [12] A. Bihlo and R. O. Popovych, “Group classification of linear evolution equations,” *Journal of Mathematical Analysis and Applications*, vol. 448, no. 2, pp. 982–1005, 2016.
- [13] Y. Liu, C. Lageman, B. D. O. Anderson, and G. Shi, “Exponential least squares solvers for linear equations over networks,” *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 2543–2548, 2017.
- [14] W. V. . Petryshyn, “On two variants of a method for the solution of linear equations with unbounded operators and their applications,” *Journal of Mathematics & Physics*, vol. 44, no. 1-4, pp. 297–312, 2016.
- [15] L. Jin, S. Li, B. Hu, M. Liu, and J. Yu, “Noise-suppressing neural algorithm for solving time-varying system of linear equations: a control-based approach,” *IEEE Transactions on Industrial Informatics*, vol. 15, no. 1, pp. 236–246, 2018.
- [16] I. E. Shparlinski, “Linear equations with rational fractions of bounded height and stochastic matrices,” *The Quarterly Journal of Mathematics*, vol. 69, no. 2, pp. 487–499, 2018.
- [17] J. Müller, “A converse result for Banach space convergence rates in Tikhonov-type convex regularization of ill-posed linear equations,” *Journal of Inverse and Ill-Posed Problems*, vol. 12, no. 5, pp. 639–646, 2017.
- [18] A. Visconti, C. V. Schiavo, and R. Peralta, “Improved upper bounds for the expected circuit complexity of dense systems of linear equations over GF(2),” *Information Processing Letters*, vol. 137, pp. 1–5, 2018.
- [19] A. K. S. Sabonchi and B. Akay, “Cryptanalysis of poly-alphabetic cipher using differential evolution algorithm,” *Tehnicki vjesnik-Technical Gazette*, vol. 27, no. 4, pp. 1101–1107, 2020.
- [20] K. Sreenu, M. Sreelatha, and W. Scheduler, “Whale optimization for task scheduling in cloud computing[[]],” *Cluster Computing*, vol. 22, no. 6, pp. 1–12, 2019.
- [21] A. Bhowmik and D. De, “mTrust: call behavioral trust predictive analytics using unsupervised learning in mobile cloud computing,” *Wireless Personal Communications*, vol. 117, no. 2, pp. 483–501, 2021.
- [22] J. P. B. Mapetu, L. Kong, and Z. Chen, “A dynamic VM consolidation approach based on load balancing using Pearson correlation in cloud computing,” *The Journal of Supercomputing*, vol. 77, no. 6, pp. 5840–5881, 2021.
- [23] V. Barthwal and M. M. S. Rauthan, “AntPu: a meta-heuristic approach for energy-efficient and SLA aware management of virtual machines in cloud computing,” *Memetic Computing*, vol. 13, no. 1, pp. 91–110, 2021.
- [24] M. A. Almaiah and A. Al-Khasawneh, “Investigating the main determinants of mobile cloud computing adoption in university campus,” *Education and Information Technologies*, vol. 25, no. 4, pp. 3087–3107, 2020.
- [25] L. S. Subhash and R. Udayakumar, “Sunflower whale optimization algorithm for resource allocation strategy in cloud computing platform,” *Wireless Personal Communications*, vol. 116, no. 4, pp. 3061–3080, 2021.
- [26] Z. Jin, “Green city economic efficiency based on cloud computing and machine learning,” *Arabian Journal of Geosciences*, vol. 14, no. 11, pp. 1–16, 2021.
- [27] A. Nr, A. Mba, and B. Ar, “Implement 3D video call using cloud computing infrastructure - ScienceDirect,” *Ain Shams Engineering Journal*, vol. 11, no. 2, pp. 363–375, 2020.
- [28] S. Rahmani, V. Khajehvand, and M. Torabian, “Burstiness-aware virtual machine placement in cloud computing systems,” *The Journal of Supercomputing*, vol. 76, no. 1, pp. 362–387, 2020.
- [29] G. Hodosi, A. Haider, and L. Rusu, “Risk factors in cloud computing relationships: a study in public organizations in Sweden,” *Procedia Computer Science*, vol. 181, no. 1, pp. 1179–1186, 2021.
- [30] S. Alebrahim and I. Ahmad, “Task scheduling for heterogeneous computing systems,” *The Journal of Supercomputing*, vol. 73, no. 6, pp. 2313–2338, 2017.
- [31] C.-Y. Chen, “Task scheduling for maximizing performance and reliability considering fault recovery in heterogeneous distributed systems,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 27, no. 2, pp. 521–532, 2016.
- [32] H. He, G. Xu, S. Pang, and Z. Zhao, “AMTS: adaptive multi-objective task scheduling strategy in cloud computing,” *China Communications*, vol. 13, no. 4, pp. 162–171, 2016.



## Research Article

# Supervision Strategy Analysis on Price Discrimination of E-Commerce Company in the Context of Big Data Based on Four-Party Evolutionary Game

Meng Xiao 

*School of Management, Shenyang University of Technology, Shenyang 110870, China*

Correspondence should be addressed to Meng Xiao; [mengxiao@sut.edu.cn](mailto:mengxiao@sut.edu.cn)

Received 30 January 2022; Revised 18 February 2022; Accepted 22 February 2022; Published 18 March 2022

Academic Editor: Daqing Gong

Copyright © 2022 Meng Xiao. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper focuses on the phenomenon of “big data killing” implied in e-commerce and discusses how to take the government as the lead to coordinately supervise the price discrimination behavior of e-commerce companies towards loyal customers. First, the four-party evolutionary game model of the government regulatory department, e-commerce platform, e-commerce company, and consumer is built. Second, the stability of the strategy choice of each game subject is analyzed. On this basis, the evolutionary stable strategy in the system based on First Law of Lyapunov is explored. Finally, the influences of key elements on system evolution are simulated and analyzed by MATLAB2021. Results demonstrate that (1) the government supervision mechanism can effectively supervise the price discrimination of e-commerce company based on big data to loyal customers; (2) when the government chooses the strict supervision strategy, reducing the information supervision cost of the e-commerce platform and the strict supervision cost of the government enable the government and the e-commerce platform to coordinate supervision and make the e-commerce company incline to choose the nondifferential pricing strategy; (3) when the government chooses the loose supervision strategy, reducing the information supervision cost of the e-commerce platform and increasing the probability of consumer discovering differential pricing and the penalties for differential pricing of e-commerce company enable the e-commerce platform and consumer to coordinate supervision, and make the e-commerce company incline to choose the nondifferential pricing strategy. The results of this study can provide theoretical guidance for the government and companies to make beneficial strategic decisions in the development of e-commerce.

## 1. Introduction

With the rise of big data, e-commerce is becoming more and more prosperous. E-commerce can bring convenience to consumers with a variety of options and also collect consumer consumption data and draw user portraits by using big data technology [1]. While the application of algorithms injects new growth drivers into social and economic development, problems caused by the unreasonable application of algorithms such as algorithm discrimination, “big data killing,” and inducing addiction also profoundly affect the normal communication in the market and destroy the market order. Online supply chain stores have different pricing based on user location. On some online booking

websites, the price of hotel rooms for Apple customer is higher than that for Windows customer. The well-known e-commerce company, Amazon, was found to use big data to “kill regular” [2]. It priced for different consumers according to their information and purchasing data on the platform. Loyal customers made purchase transactions based on their trust and path dependence on the Amazon platform, but due to the asymmetry of information in the transaction process, some regulars pay higher prices than strangers. This “big data killing” behavior has exposed the hidden dangers of moral hazard in the e-commerce market and makes the industry encounter an unprecedented crisis of trust. “Big data killing” has become an urgent problem to be solved in the fast development of online business [3].



The essence of big data killing is price discrimination. Price discrimination refers to formulating different price strategies for different customer groups. However, in traditional business, both “stranger” and “regular” may be discriminated against, while with the participation of algorithm technology, there are more “killing regular” in Internet business. Even in the process of “killing regular,” big data has become a necessary tool. Each platform will collect a lot of user information, and then the company uses technology to offer different prices and discounts for different customers based on the information. Traditionally, companies have not been able to predict the upper limit of the price that buyers want to pay, but based on the technology of big data, the companies can determine the maximum willing price with a high degree of accuracy with sophisticated algorithms [4]. As the collection of consumer data becomes more common, online companies are now more capable of price discrimination than ever [5]. As customer of the Internet commercial company, VIP customer with higher loyalty and stronger consumption power pay much more for the same service than new customers, but gain even lower service quality. Big data killing will cause a variety of harm. Moriarty [6] proposed that customer information is widely used in online retail pricing, and although the benefit of online retailers will increase, price discrimination can cause serious fairness concerns and even violation of regulations and laws. Antimonopoly issues in the digital economy, especially the antimonopoly issues of big data and discrimination algorithms, have been brought to the attention of experts and practitioners. “Killing regular” is algorithmic price discrimination, with which online platforms charge long-term customers higher prices. It is believed that this kind of price discrimination violates the law on antimonopoly and should be held accountable according to the relevant law. The Cyberspace Administration of China (CAC) issued the regulations on The Management of Algorithm Recommendation for Internet Information Services to regulate the “big data killing,” stepping into the era of strict supervision of the industry related to algorithm recommendation. The EU also prohibits discrimination on certain grounds and strictly regulates unfair business practices in B2C relationships [7].

Although some studies have carried out related discussion on the problem of big data killing [2, 8], some solutions are proposed [9, 10]. However, existing studies are mostly limited to the pricing between e-commerce companies and consumers [11, 12], the strategic choice between e-commerce platform and consumer, and the supervision strategy choice between e-commerce platform and government [9, 13]. There are few systematic studies on the four-party strategy composed of multiple subjects related to “big data killing.” Therefore, this study establishes an evolutionary game model dominated by government supervision that affects the decision-making of consumer, e-commerce company, and platform, analyzes and simulates that different supervision costs of government and e-commerce platform, consumer discovery levels, and the penalties for differential pricing of e-commerce company affect

system equilibrium, evolutionarily stable strategy, and the pricing strategy of e-commerce company, and also establishes the platform-consumer-government collaborative supervision mechanism for e-commerce company pricing behavior. This research contributes to curbing the “big data killing” behavior of e-commerce company, enhance consumers’ confidence in online shopping, and has a positive effect in promoting the development of e-commerce.

## 2. Related Literature

Existing research on price discrimination in e-commerce companies mainly focuses on three aspects: the prevalence of price discrimination by using customer information, the influence factors of price discrimination, and the supervision and management of price discrimination:

On the prevalence of price discrimination by e-commerce companies using customer information, although many media outlets provide various evidence of price discrimination, most of them are not based on scientific and systematic methods. Therefore, scholars have researched whether e-commerce companies use big data to discriminate against consumers in price. Botta and Klaus [14] qualitatively proposed that algorithmic price discrimination is different from offline differential pricing and is related to the collection of consumer information, which is a unique feature of the digital economy. With the wide application of big data and the gradual deepening of algorithm technology, the e-commerce company can price discriminate against consumers with great precision [4], and these were confirmed empirically [4, 15]. The pricing ecosystem of the online platform is a dynamic pricing system [15]. Algorithmic price discrimination [16], artificial intelligence techniques, and digital system fingerprints [15] enable the e-commerce company to have the ability of price discrimination. Price discrimination is not only widespread in the field of commodity sales, and there are also discrimination and price difference by using customer information in the field of online car-hailing [8] and the field of advertising recommendation [7, 17]. While consumers benefit from accurate recommendations, sellers may use this information to discriminate on price. Thus, price discrimination is not favored by people [18].

Scholars have done a lot of studies on the influence factors of price discrimination in e-commerce companies. Some scholars believe that the premise of “big data killing” is the information asymmetry between e-commerce company and customers [1]. Consumer information data is an influencing factor for the e-commerce company to be able to discriminate in price [19], such as consumer characteristics, location [14], etc., and these data also relate to consumers’ privacy [12]. Nuccio and Marco [20] studied how pricing technology and information transparency are changing merchants’ pricing behavior in online transactions. The price sensitivity and heterogeneity of consumers are factors that affect e-commerce company to set price differentials [11]. Some scholars have analyzed the effects of reference price and search cost on differential pricing and find that

consumers' search cost has become one of the obstacles affecting consumers' online shopping, which has formed an unequal situation for consumers [21] and has become a tool for e-commerce companies to formulate differential prices [22, 23]. The target of "big data killing" of e-commerce companies is focused on loyal consumers, which has been confirmed by many scholars. For example, Tang et al. [24] found in the research on the group-buying market that with the improvement of consumer retention rate, the best strategy of sellers is changed from quality difference to price discrimination. Chandra and Lederman [25] argued that if consumers have differences in potential willingness to pay and brand loyalty, e-commerce companies may increase price differences among some consumers while reducing price differences among the other consumers. Although differential pricing is an important way for e-commerce companies to obtain profits [26], its focus on loyal consumers is contrary to the principle of fair pricing [24], which will reduce consumer satisfaction and create distrust [27, 28].

After the problem of "big data killing" was exposed, it has been attracted widespread attention by scholars, and its supervision and governance have also become an important research topic. Bar-Gill [29] proposed that the normative evaluation of price discrimination depends on the object of discrimination, and the algorithmic price discrimination has the advantages to improve efficiency, but it will harm consumers, which should be governed by rules set by regulators to seriously exploit the potential of personalization. Yu and Li [9] also believed that consumers' discovery and reporting of being "killed" is the mean to monitor price discrimination of e-commerce company. Xing et al. [3] found that when regulars account for a high proportion of platform customer, giving consumers the right to data portability can curb the phenomenon of "big data killing" to a certain extent. In addition to consumers' self-discovery of price discrimination by the e-commerce company, many scholars believe that with the help of government supervision [13], increasing penalties and the commission coefficient of government departments [30] can effectively reduce the "killing regular" pricing tendency of e-commerce platforms. However, in the supervision process of existing research, there was little distinction between e-commerce platform and company, and the research is carried out in a mixed way. Most of the discussions focus on the pricing of e-commerce platforms known for its scale. Differential pricing of e-commerce company on the platform is rarely discussed separately, and there is still a lack of research considering multichannel collaborative supervision.

Existing studies have adopted a variety of methods for the problem of price discrimination in the e-commerce company. For example, the dynamic pricing method is used in specific pricing. Lindgren et al. [31] studied dynamic pricing by intertemporal price discrimination theory and proposed that retailers should change prices randomly over time. Chevalier and Kashyap [32] proposed the method for aggregating prices when retailers use periodic sales to discriminate price against heterogeneous customers. Tremblay

[5] designed more efficient Pareto price discrimination. Game methods are often used in the selection of pricing strategies. Choe et al. [33] analyzed pricing strategy with a two-stage dynamic game model. Zhou et al. [34] adopted two-stage game analysis on joint pricing and bandwidth demand optimization. On the game of price discrimination, the bounded rationality assumption in the evolutionary game makes the research more realistic [30], so many scholars use evolutionary game methods to study this problem [1, 13, 30, 35] and extended to multiple fields of online transactions, such as manufacturing business [36]. However, most studies are limited in the two-party game [22, 37], it is still unclear to analyze the relationship and role of e-commerce company, consumer, e-commerce platform, and the government in the "big data killing" problem system, and their decision-making mechanism needs further research.

Therefore, as the price discrimination of e-commerce companies is generated with new technologies, the existing research on this phenomenon is still in the exploratory stage. Most perspectives of the previous research are from both sides of the transaction in traditional business, and there are few differences in the analysis of the e-commerce platform and the companies in the platform. Moreover, the supervision on the differential pricing of the e-commerce company using big data technology to the loyal customers is not very perfect, and some policies and supervision methods are still under discussion. This study systematically analyzes the government, e-commerce platform, e-commerce company, and consumer involved in the supervision of "big data killing," which makes up for the insufficiency of the existing research and provides useful help for further regulating such behavior.

### 3. Materials and Methods

*3.1. Problem Description.* The e-commerce company will use the platform to collect consumer information during the operation in the network platform. Based on the information provided by the platform, e-commerce company analyzes consumers and raise prices by judging their consumption habits. The pricing strategy of "big data killing" is price discrimination caused by e-commerce company using the feature of opaque information in the online transaction process to different pricing of consumers through big data and complex algorithms. This kind of behavior will bring consumers' distrust of e-commerce companies and e-commerce platforms, which is not conducive to the development of e-commerce. Therefore, both the government regulatory department and e-commerce platforms should take necessary measures to supervise the price discrimination behavior of e-commerce companies. This study mainly discusses the following three questions: (1) in the context of big data development, how can the government regulatory department take supervision measures to reduce the proportion of price discrimination by e-commerce company? (2) How can e-commerce platform be motivated to supervise information on e-commerce companies? (3) How can

consumers be guided to actively safeguard their rights and interests and maintain consumption fairness.

This study builds a multi-agent game model for the supervision of price discrimination in e-commerce companies involving the e-commerce platform, the e-commerce company, the consumer, and the government regulatory department. The logical relationship among four-party game subjects is shown in Figure 1.

**3.2. Model Assumption.** To build the multisubject supervision model of the e-commerce company pricing in the background of big data, the behavioral strategies of government regulatory department, e-commerce platform, e-commerce company, and consumer are studied, and the following assumptions are made.

*Assumption 1.* Government regulatory department, e-commerce platform, the e-commerce company, and consumer are selected as the game subjects. Each game subject is bounded rationality and pursues the maximization of their interests in e-commerce transactions. Due to the information asymmetry between game subjects, random behavior strategies, and interactive effects, the optimal strategy cannot be obtained through one game. It is necessary to continuously try and learn in multiple rounds of games to improve the strategy, to formulate the best match of behavioral decision. Therefore, the evolutionary game should be used to analyze the four-party equilibrium strategy. The proportion of e-commerce company implementing nondifferential pricing is represented as  $x$  ( $0 \leq x \leq 1$ ), and the proportion of e-commerce company implementing differential pricing is denoted as  $(1-x)$ ; the proportion of consumer loyalty is represented as  $y$  ( $0 \leq y \leq 1$ ) and the proportion of consumer disloyalty is represented as  $(1-y)$ ; the proportion of e-commerce platform to supervise company information is represented as  $z$  ( $0 \leq z \leq 1$ ), and the proportion of e-commerce platform with information nonsupervision is denoted as  $(1-z)$ ; the proportion of the government regulatory department strictly supervising e-commerce platform and company is denoted as  $r$  ( $0 \leq r \leq 1$ ), and the proportion of loosely supervises e-commerce platform and the company is denoted as  $(1-r)$ .

*Assumption 2.* The benefit of nondifferential pricing of the e-commerce company is  $P_n$ , and the basic benefit of differential pricing is  $P_d$ . When the e-commerce company implements differential pricing for loyal consumer, additional benefit  $\Delta P$  can be obtained due to the increase in selling price, and  $P_d < P_n < P_d + \Delta P$ . The probability of loyal consumers discovering differential pricing of the e-commerce company is  $\alpha$ . When consumer purchases goods, the utility obtained by the loyal consumer is  $U_b$ , and the utility obtained by the disloyal consumer is  $U_d$ , and  $U_l > U_d$ . The reputation value of the loyal consumer to the e-commerce company is  $T_e$  and the reputation value of the loyal consumer to the e-commerce platform is  $T_p$ .

*Assumption 3.* When the government strictly supervises, if price discrimination of the e-commerce company is found, loyal consumers who are subject to differential pricing will be compensated with the compensation amount of  $M$ ; When the government loosely supervises, if the loyal consumer is the price-sensitive consumer, he may use Internet information for comparison and analysis, and then find that he has been "killed". If the cost of reporting is small and the procedure is simple, the consumer will carry out to inform the government regulatory department, and then the e-commerce company must be forced to compensate the consumer. The consumer's complaint cost is  $C_c$ .

*Assumption 4.* The normal benefit that the government obtains from the operation of the e-commerce platform is  $S$ . The cost of strict supervision by government departments is  $C_g$ . The social benefit obtained by the government is  $R$  if there is no price discrimination by the e-commerce company. If the government adopts the loose supervision policy, consumer complaints will bring social reputation loss as  $N$ . After receiving the information, the e-commerce company for price discrimination will be penalized by the government regulatory department, and the fine will be  $I_e$ .

*Assumption 5.* The price discrimination of e-commerce company depends on the information provided by the platform. The benefit of the platform reasonably providing information to the e-commerce company is  $W$ , and the cost of the platform information supervision on e-commerce company is  $C_p$ . When the e-commerce platform finds the price discrimination of e-commerce company on the consumer, the fine to e-commerce company is  $F$ . The e-commerce platform and consumers share this fine in the ratio of  $\beta$  and  $1-\beta$ . When the government finds price discrimination by the e-commerce company, it will impose the fine of  $I_p$  for the platform's unfavorable supervision to e-commerce company information.

The parameters are described in Table 1.

**3.3. Model Framework.** According to the above analysis, the mixed-strategy game matrix of the four-party game subjects of government regulatory department, e-commerce platform, e-commerce company, and consumer is shown in Table 2.

### 3.4. Model Analysis

**3.4.1. Strategy Stability Analysis of the E-Commerce Company.** Assuming that the expected benefit of the e-commerce company when choosing the nondifferential pricing strategy is  $U_{11}$ , the expected benefit of the e-commerce company when choosing the differential pricing strategy is  $U_{12}$ , and the average expected benefit of the e-commerce company is  $\bar{U}_1$ , which are defined as follows:

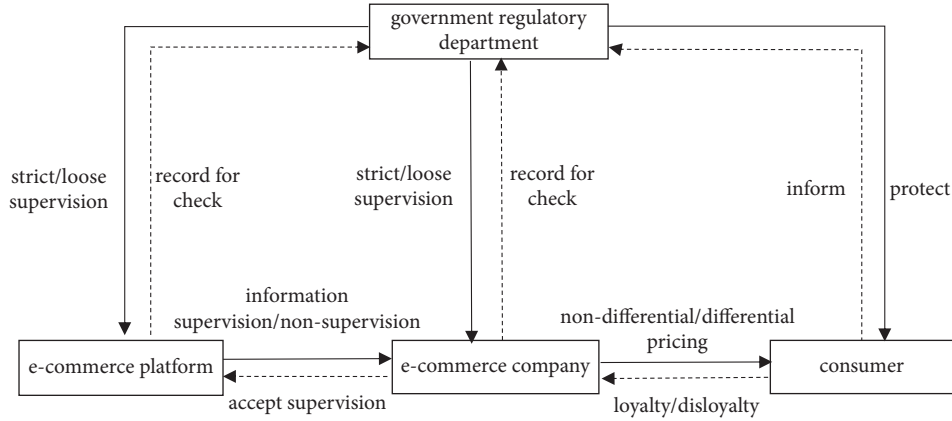


FIGURE 1: Game model logic relationship of multisubject supervision on e-commerce company pricing.

TABLE 1: Parameter description.

Parameter	Description
$P_n$	The benefit of nondifferential pricing by e-commerce company to consumer
$P_d$	The benefit of differential pricing by e-commerce company to consumer
$\Delta P$	The additional benefit of differential pricing by e-commerce company to the loyal consumer
$M$	Compensation of e-commerce company to the loyal consumer for differential pricing
$T_e$	The reputation value of the loyal consumer to e-commerce company
$U_l$	The utility obtained by the loyal consumer from purchasing goods
$U_d$	The utility obtained by the disloyal consumer from purchasing goods
$C_c$	The cost of consumer complaint
$\alpha$	Probability of loyal consumer discovering differential pricing under government loose supervision, and $\alpha \in [0, 1]$
$C_g$	The cost of strict supervision by the government regulatory department
$N$	Social reputation loss caused by differential pricing under government loose supervision
$R$	The social benefit of nondifferential pricing under the government strict supervision
$I_e$	Fine by government regulatory department for differential pricing to e-commerce company
$I_p$	Fine by government imposed on the platform for nonsupervision of e-commerce company information resulting in differential pricing
$S$	The normal benefit obtained by the government from the operation of the e-commerce platform
$W$	The benefit of the platform reasonably providing information to the e-commerce company
$C_p$	The cost of the platform's information supervision on the e-commerce company
$F$	Fines imposed by the platform to e-commerce company for differential pricing during information supervision
$\beta$	The proportion of the fine imposed by the e-commerce platform for differential pricing of e-commerce company, $\beta \in [0, 1]$
$T_p$	The reputation value of the loyal consumer to the e-commerce platform

TABLE 2: Game model benefit matrix of government regulatory department, e-commerce platform, e-commerce company, and consumer

Strategy choice		Government regulatory department				
		E-commerce company	Strict supervision, $r$		Loose supervision, $1 - r$	
			Loyalty $y$	Disloyalty $1 - y$	Loyalty $y$	Disloyalty $1 - y$
E-commerce platform	Information supervision $z$	Nondifferential pricing $x$	$P_n + T_e$	$P_n$	$P_n + T_e$	$P_n$
			$U_l$	$U_d$	$U_l$	$U_d$
	Information nonsupervision $1 - z$	Differential pricing $1 - x$	$W - C_p + T_p$	$W - C_p$	$W - C_p + T_p$	$W - C_p$
			$S - C_g + R$	$S - C_g + R$	$S$	$S$
E-commerce company	Information supervision $z$	Nondifferential pricing $x$	$P_d + \Delta P + T_e - M - I_e - F$	$P_d - I_e - F$	$P_d + \Delta P + T_e - \alpha M - \alpha I_e - F$	$P_d - F$
			$U_l - \Delta P + M + (1 - \beta)F$	$U_d$	$U_l - \Delta P - C_c + \alpha M + (1 - \beta)F$	$U_d$
	Information nonsupervision $1 - z$	Differential pricing $1 - x$	$W - C_p + \beta F + T_p$	$W - C_p + F$	$W - C_p + \beta F + T_p$	$W - C_p + F$
			$S - C_g + I_e$	$S - C_g + I_e$	$S + \alpha I_e - N$	$S - N$
Government regulatory department	Information supervision $z$	Nondifferential pricing $x$	$P_n + T_e$	$P_n$	$P_n + T_e$	$P_n$
			$U_l$	$U_d$	$U_l$	$U_d$
	Information nonsupervision $1 - z$	Differential pricing $1 - x$	$W + T_p$	$W$	$W + T_p$	$W$
			$S - C_g + R$	$S - C_g + R$	$S$	$S$
Consumer	Information supervision $z$	Nondifferential pricing $x$	$P_d + \Delta P + T_e - M - I_e$	$P_d - I_e$	$P_d + \Delta P + T_e - \alpha M - \alpha I_e$	$P_d$
			$U_l - \Delta P + M$	$U_d$	$U_l - \Delta P - C_c + \alpha M$	$U_d$
	Information nonsupervision $1 - z$	Differential pricing $1 - x$	$W - I_p + T_p$	$W - I_p$	$W - \alpha I_p + T_p$	$W$
			$S - C_g + I_e + I_p$	$S - C_g + I_e + I_p$	$S + \alpha I_e + \alpha I_p - N$	$S - N$

$$\begin{aligned}
U_{11} &= yzr(P_n + T_e) + (1 - y)zrP_n + y(1 - z)r(P_n + T_e) + (1 - y)(1 - z)rP_n \\
&\quad + yz(1 - r)(P_n + T_e) + (1 - y)z(1 - r)P_n + y(1 - z)(1 - r)(P_n + T_e) + (1 - y)(1 - z)(1 - r)P_n \\
&= P_n + yT_e, \\
U_{12} &= yzr(P_d + \Delta P + T_e - M - I_e - F) + (1 - y)zr(P_d - I_e - F) \\
&\quad + y(1 - z)r(P_d + \Delta P + T_e - M - I_e) + (1 - y)(1 - z)r(P_d - I_e) \\
&\quad + yz(1 - r)(P_d + \Delta P + T_e - \alpha M - \alpha I_e - F) + (1 - y)z(1 - r)(P_d - F) \\
&\quad + y(1 - z)(1 - r)(P_d + \Delta P + T_e - \alpha M - \alpha I_e) + (1 - y)(1 - z)(1 - r)P_d \\
&= P_d + y(\Delta P + T_e) - y(M + I_e)[r + (1 - r)\alpha] - (1 - y)rI_e - zF, \\
\bar{U}_1 &= xU_{11} + (1 - x)U_{12}.
\end{aligned} \tag{1}$$

According to the Malthusian dynamic equation, the replication dynamic equation of the e-commerce company is obtained as follows:

$$\begin{aligned}
F(x) &= \frac{dx}{dt} \\
&= x(U_{11} - \bar{U}_1) \\
&= x(1 - x)\{P_n - P_d - y\Delta P + y(M + I_e) \\
&\quad \cdot [r + (1 - r)\alpha] + (1 - y)rI_e + zF\}.
\end{aligned} \tag{2}$$

The first partial derivative of  $F(x)$  for  $x$  is as follows:

$$\begin{aligned}
F'_x(x) &= (1 - 2x)\{P_n - P_d - y\Delta P + y(M + I_e) \\
&\quad \cdot [r + (1 - r)\alpha] + (1 - y)rI_e + zF\}.
\end{aligned} \tag{3}$$

Based on the stability theorem of differential equations, the e-commerce company implements the strategy of nondifferential pricing in the stable state must meet the conditions:  $F(x) = 0$ , and  $F'_x(x) < 0$ .

**Proposition 1.** *When  $r > r_0$ , the stable strategy of the e-commerce company is nondifferential pricing; when  $r < r_0$ , the stable strategy of the e-commerce company is differential pricing; when  $r = r_0$ , the e-commerce company cannot determine the stable strategy. Where the threshold is as follows:*

$$r_0 = \frac{P_d + y\Delta P - P_n - \alpha y(M + I_e) - zF}{(1 - \alpha)yM + (1 - \alpha y)I_e}. \tag{4}$$

*Proof.* Assume  $H(r) = P_n - P_d - y\Delta P + y(M + I_e)[r + (1 - r)\alpha] + (1 - y)rI_e + zF$ , when  $y[M - \alpha(M + I_e)] > 0$ ,  $(\partial H/\partial r) > 0$ , then  $H(r)$  is considered to be an increasing function of  $r$ . When  $r > r_0$ ,  $H(r) > 0$ ,  $F(x)|_{x=1} = 0$ , and  $F'_x(x)|_{x=1} < 0$ , so  $x=1$  has stability; When  $r < r_0$ ,  $H(r) < 0$ ,  $F(x)|_{x=0} = 0$ , and  $F'_x(x)|_{x=0} < 0$ , so  $x=0$  has stability; when  $r = r_0$ ,  $H(r) = 0$ ,  $F(x) = 0$ , and  $F'_x(x) = 0$ , so  $x$  is stable at all levels in the range of 0 to 1, that is, the company's strategy does not change over time, regardless of the proportion of company choosing to price differentially.

Proposition 1 states that the increase of the proportion of the government strict supervision to e-commerce company will change the stable strategy of e-commerce company from

differential pricing to nondifferential pricing; Similarly, the decline of the proportion of the government strict supervision to e-commerce company will change the stable strategy of e-commerce company from nondifferential pricing to differential pricing. Therefore, the government's strict supervision for e-commerce company is essential, and the government should take measures to improve strict supervision for the e-commerce company.

Based on Proposition 1, the phase diagram of the strategy evolution of e-commerce company is shown in Figure 2.

Inference 1: with the increase of the value of  $P_n$ ,  $M$ ,  $I_e$ ,  $F$ , and  $\alpha$ , the e-commerce company is more inclined to implement the nondifferential pricing strategy, when other parameters remain unchanged. Similarly, with the increase of the value of  $P_d$  and  $\Delta P$ , the e-commerce company is more inclined to implement the differential pricing strategy. It shows that the proportion of e-commerce company implementing nondifferential pricing strategy is directly proportional to the benefits of nondifferential pricing, the fines imposed by the government and platform on e-commerce company for differential pricing and the probability of consumers' discovery, and inversely proportional to the benefits of e-commerce company implementing differential pricing strategy.  $\square$

*Proof.* Since  $r_0 = (P_d + y\Delta P - P_n - \alpha y(M + I_e) - zF) / ((1 - \alpha)yM + (1 - \alpha y)I_e)$ , the volume of  $V_{x1}$  in Figure 2 represents the proportion of nondifferential pricing by the e-commerce company, and the corresponding volume of  $V_{x0}$  represents the proportion of differential pricing by the e-commerce company. When the value of  $P_n$ ,  $M$ ,  $I_e$ ,  $F$ , and  $\alpha$  gradually increases, the value of  $r_0$  will gradually decrease, and the volume of  $V_{x1}$  will increase at this time, indicating that the proportion of e-commerce company to implement nondifferential pricing increases; When the value of  $P_d$  and  $\Delta P$  gradually increases, the value of  $r_0$  will gradually increase, and the volume of  $V_{x1}$  will decrease at this time, indicating that the proportion of e-commerce company to implement nondifferential pricing decreases.  $\square$

**3.4.2. Strategy Stability Analysis of the Consumer.** Assuming that the expected benefit of the consumer when choosing loyalty strategy to e-commerce company is  $U_{21}$ , the expected benefit of the consumer when choosing disloyalty

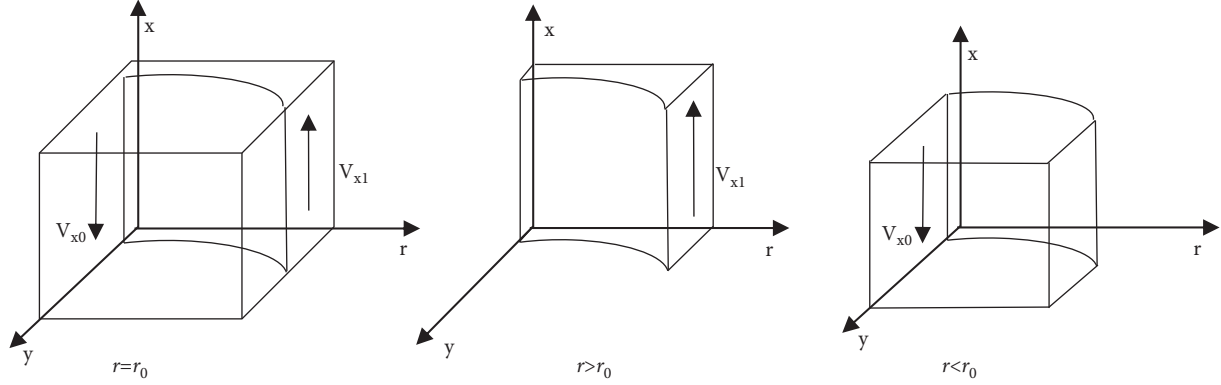


FIGURE 2: Phase diagram of strategy evolution of e-commerce company.

strategy to e-commerce company is  $U_{22}$ , and the average expected benefit of the consumer is  $\overline{U}_2$ , which are defined as follows:

$$\begin{aligned}
 U_{21} &= xzrU_l + (1-x)zr[U_l - \Delta P + M + (1-\beta)F] \\
 &\quad + x(1-z)rU_l + (1-x)(1-z)r(U_l - \Delta P + M) \\
 &\quad + xz(1-r)U_l + (1-x)z(1-r) \\
 &\quad \cdot [U_l - \Delta P - C_c + \alpha M + (1-\beta)F] \\
 &\quad + x(1-z)(1-r)U_l + (1-x)(1-z) \\
 &\quad \cdot (1-r)(U_l - \Delta P - C_c + \alpha M) \\
 &= U_l - (1-x)\Delta P + (1-x)[M + (1-r)\alpha M] \\
 &\quad - (1-x)(1-r)C_c \\
 &\quad + (1-x)z(1-\beta)F, \\
 U_{22} &= xzrU_d + (1-x)zrU_d \\
 &\quad + x(1-z)rU_d \\
 &\quad + (1-x)(1-z)rU_d \\
 &\quad + xz(1-r)U_d \\
 &\quad + (1-x)z(1-r)U_d \\
 &\quad + x(1-z)(1-r)U_d \\
 &\quad + (1-x)(1-z)(1-r)U_d \\
 &= U_d, \\
 \overline{U}_2 &= yU_{21} + (1-y)U_{22}.
 \end{aligned} \tag{5}$$

According to the Malthusian dynamic equation, the replication dynamic equation of consumer is obtained as follows:

$$\begin{aligned}
 F(y) &= \frac{dy}{dt} \\
 &= y(U_{21} - \overline{U}_2) \\
 &= y(1-y) \left\{ \begin{array}{l} U_l - U_d - (1-x)\Delta P \\ + (1-x)[M + (1-r)\alpha M] \\ - (1-x)(1-r)C_c + (1-x)z(1-\beta)F \end{array} \right\}.
 \end{aligned} \tag{6}$$

The first partial derivative of  $F(y)$  for  $y$  is as follows:

$$\begin{aligned}
 F'_y(y) &= (1-2y)\{U_l - U_d - (1-x)\Delta P + (1-x)[M + (1-r)\alpha M] \\
 &\quad - (1-x)(1-r)C_c + (1-x)z(1-\beta)F\}.
 \end{aligned} \tag{7}$$

Based on the stability theorem of differential equations, consumer implements the strategy of loyalty in the stable state must meet the conditions:  $F(y) = 0$ , and  $F'_y(y) < 0$ .

**Proposition 2.** When  $x > x_0$ , the stable strategy of the consumer is loyalty; when  $x < x_0$ , the stable strategy of the consumer is disloyalty; when  $x = x_0$ , the consumer cannot determine the stable strategy. Where the threshold is as follows:

$$x_0 = \frac{U_l - U_d + M - \Delta P + (1-r)(\alpha M - C_c) + z(1-\beta)F}{M - \Delta P + (1-r)(\alpha M - C_c) + z(1-\beta)F}. \tag{8}$$

*Proof.* Assume  $H(x) = U_l - U_d - (1-x)\Delta P + (1-x)[M + (1-r)\alpha M] - (1-x)(1-r)C_c + (1-x)z(1-\beta)F$ , when  $M - \Delta P + (1-r)(\alpha M - C_c) + z(1-\beta)F > 0$ ,  $(\partial H/\partial x) > 0$ ,  $H(x)$  is considered to be an increasing function of  $x$ . When  $x > x_0$ ,  $H(x) > 0$ ,  $F(y)|_{y=1} = 0$ , and  $F'_y(y)|_{y=1} < 0$ , so  $y = 1$  has stability; When  $x < x_0$ ,  $H(x) < 0$ ,  $F(y)|_{y=0} = 0$ , and  $F'_y(y)|_{y=0} < 0$ , so  $y = 0$  has stability; When  $x = x_0$ ,  $H(x) = 0$ ,  $F(y) = 0$ , and  $F'_y(y) = 0$ , so  $y$  is stable at all levels in the range of 0 to 1, that is, the consumer's strategy does not change over time, regardless of the proportion of consumer choosing to be loyal.

Proposition 2 states that the increase of the proportion of nondifferential pricing of e-commerce company will change the stable strategy of consumer from disloyalty to loyalty; Similarly, the decline of the proportion of nondifferential pricing of e-commerce company will change the stable strategy of consumer from loyalty to disloyalty. Therefore, e-commerce company should reduce the degree of difference in pricing for consumers and try to retain consumers.

Based on Proposition 2, the phase diagram of the strategy evolution of consumer is shown in Figure 3.

Inference 2: with the increase of the value of  $U_l$ ,  $M$ ,  $F$ ,  $\alpha$ , and  $\beta$ , the consumer is more inclined to be loyalty strategy to



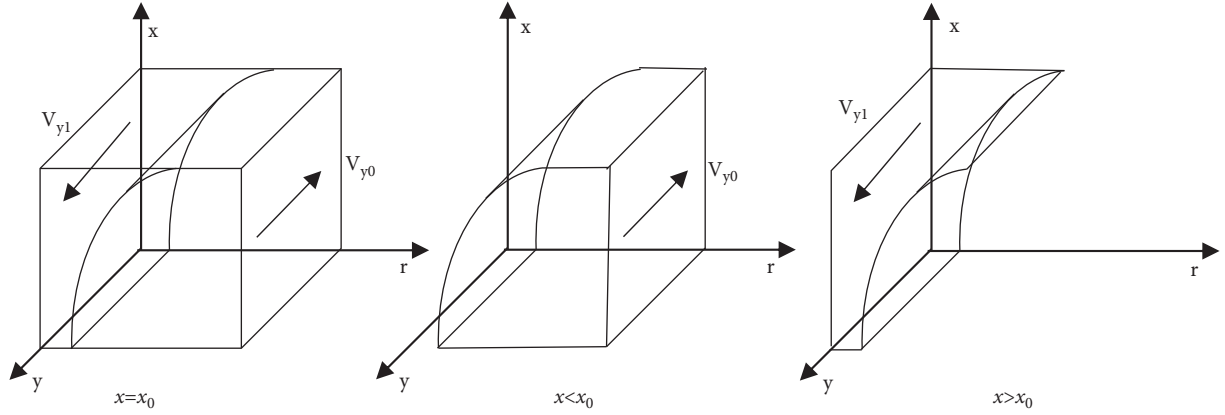


FIGURE 3: Phase diagram of strategy evolution of consumer.

the e-commerce company, when other parameters remain unchanged. Similarly, with the increase of the value of  $U_d$ ,  $\Delta P$ , and  $C_c$ , the consumer is more inclined to be disloyalty strategy to the e-commerce company. It shows that the proportion of consumer being loyalty strategy to e-commerce company is directly proportional to the utility obtained by the loyal consumer from purchasing goods, the fines imposed by the government and e-commerce platform for differential pricing of e-commerce company, and the probability of consumers' discovery, and inversely proportional to the utility obtained by the disloyal consumer in purchasing goods, the additional benefit obtained by the e-commerce company in implementing differential pricing, the proportion of fines imposed by the platform to the e-commerce company and the cost of consumer complaints.  $\square$

*Proof.* Since  $x_0 = 1 - (U_l - U_d / \Delta P + (1-r)C_c - [1 + (1-r)\alpha]M - z(1-\beta)F)$ , the volume of  $V_{y1}$  in Figure 3 represents the proportion of loyalty to e-commerce company by the consumer, and the corresponding volume of  $V_{y0}$  represents the proportion of disloyalty to e-commerce company by the consumer. When the value of  $U_b$ ,  $M$ ,  $I_e$ ,  $F$ , and  $\alpha$  gradually increases, the value of  $x_0$  will gradually decrease, and the volume of  $V_{y1}$  will increase at this time, indicating that the proportion of loyalty to e-commerce company by the consumer increases; When the value of  $U_d$ ,  $\Delta P$ ,  $\beta$  and  $C_c$  gradually increase, the value of  $x_0$  will gradually increase, and the volume of  $V_{y1}$  will decrease at this time, indicating that the proportion of loyalty to e-commerce company by consumer decreases.  $\square$

### 3.4.3. Strategy Stability Analysis of E-Commerce Platform.

Assuming that the expected benefit of the e-commerce platform when choosing the information supervision strategy is  $U_{31}$ , the expected benefit of the e-commerce platform when choosing the information nonsupervision strategy is  $U_{32}$ , and the average expected benefit of the e-commerce platform is  $\bar{U}_3$ , which are defined as follows:

$$\begin{aligned}
 U_{31} &= xy r(W - C_p + T_p) + x(1-y)r(W - C_p) \\
 &\quad + (1-x)(yrW - C_p + T_p + \beta F) \\
 &\quad + (1-x)(1-y)r(W - C_p + T_p) \\
 &\quad + xy(1-r)(W - C_p + \beta F) \\
 &\quad + x(1-y)(1-r)(W - C_p) \\
 &\quad + (1-x)y(1-r)(W - C_p + T_p + \beta F) \\
 &\quad + (1-x)(1-y)(1-r)(W - C_p + \beta F) \\
 &= W - C_p + F + yT_p - xF, \\
 U_{32} &= xy r(W + T_p) + x(1-y)rW \\
 &\quad + (1-x)yr(W - I_p + T_p) \\
 &\quad + (1-x)(1-y)r(W - I_p) \\
 &\quad + xy(1-r)(W + T_p) \\
 &\quad + x(1-y)(1-r)W \\
 &\quad + (1-x)y(1-r)(W - \alpha I_p + T_p) \\
 &\quad + (1-x)(1-y)(1-r)W \\
 &= W + yT_p - (1-x)rI_p - (1-x)y(1-r)I_p, \\
 \bar{U}_3 &= zU_{31} + (1-z)U_{32}.
 \end{aligned} \tag{9}$$

According to the Malthusian dynamic equation, the replication dynamic equation of e-commerce platform is obtained as follows:

$$\begin{aligned}
 F(z) &= \frac{dz}{dt} \\
 &= z(U_{31} - \bar{U}_3) \\
 &= z(1-z)\{\beta F - C_p + yT_p - x\beta F \\
 &\quad - (1-x)rI_p - (1-x)y(1-r)I_p\}.
 \end{aligned} \tag{10}$$

The first partial derivative of  $F(z)$  for  $z$  is as follows:

$$\begin{aligned}
 F'_z(z) &= (1-2z)\{\beta F - C_p + yT_p - x\beta F \\
 &\quad - (1-x)rI_p - (1-x)y(1-r)I_p\}.
 \end{aligned} \tag{11}$$

Based on the stability theorem of differential equations, e-commerce platform implements the strategy of information supervision in the stable state must meet the conditions:  $F(z) = 0$ , and  $F'_z(z) < 0$ .

**Proposition 3.** *When  $y > y_0$ , the e-commerce platform will choose information supervision as the stable strategy; when  $y < y_0$ , the e-commerce platform will choose information nonsupervision as the stable strategy; when  $y = y_0$ , the e-commerce platform cannot determine the stable strategy. Where the threshold is as follows:*

$$y_0 = \frac{C_p + x\beta F + (1-x)rI_p - \beta F}{T_p - (1-x)(1-r)I_p}. \quad (12)$$

*Proof.* Assume  $H(y) = F - C_p + yT_p - xF - (1-x)y(1-r)I_p$ , when  $T_p - (1-x)(1-r)I_p > 0$ ,  $(\partial H/\partial x) > 0$ ,  $H(y)$  is considered to be an increasing function of  $y$ . When  $y > y_0$ ,  $H(y) > 0$ ,  $F(z)|_{z=1} = 0$ , and  $F'_z(z)|_{z=1} < 0$ , so  $z = 1$  has stability; When  $y < y_0$ ,  $H(y) < 0$ ,  $F(z)|_{z=0} = 0$ , and  $F'_z(z)|_{z=0} < 0$ , so  $z = 0$  has stability; When  $z = z_0$ ,  $H(y) = 0$ ,  $F(z) = 0$ , and  $F'_z(z) = 0$ , so  $z$  is stable at all levels in the range of 0 to 1, that is, the e-commerce platform's strategy does not change over time, regardless of the proportion of e-commerce platform choosing information supervision.

Proposition 3 states that the increase of the proportion of consumer loyalty will change the stable strategy of e-commerce platform from information nonsupervision to information supervision. Similarly, the decline of the proportion of consumer loyalty will change the stable strategy of e-commerce platform from information supervision to information nonsupervision. Therefore, if the consumer can be loyal to the e-commerce company in the platform, the platform will also actively supervise its subordinate company.

Based on Proposition 3, the phase diagram of the strategy evolution of the e-commerce platform is shown in Figure 4.

*Inference 3:* with the increase of the value of  $F$ ,  $\beta$ , and  $T_p$ , the e-commerce platform is more inclined to implement the information supervision strategy, when other parameters remain unchanged. Similarly, with the increase of the value of  $C_p$  and  $I_p$ , the e-commerce platform is more inclined to implement the information nonsupervision strategy. It shows that the proportion of e-commerce platform implementing information supervision strategy is directly proportional to the fines imposed by the platform for differential pricing of e-commerce company, the proportion of fines imposed by the e-commerce platform for differential pricing of e-commerce company, and the reputation value brought by the loyal consumer to the platform, and inversely proportional to the cost of the platform's information supervision on e-commerce company and the fines by government imposed on the platform for nonsupervision of e-commerce company information resulting in differential pricing.  $\square$

*Proof.* Since  $y_0 = (C_p + (1-x)rI_p - (1-x)\beta F) / (T_p - (1-x)(1-r)I_p)$ , the volume of  $V_{z1}$  in Figure 4 represents the

proportion of information supervision of e-commerce company by the platform, and the corresponding volume of  $V_{z0}$  represents the proportion of information non-supervision by the platform. When the value of  $F$ ,  $\beta$ , and  $T_p$  gradually increase, the value of  $y_0$  will gradually decrease, and the volume of  $V_{z1}$  will increase at this time, indicating that the proportion of e-commerce platform to implement information supervision increases; When the value of  $C_p$  and  $I_p$  gradually increases, the value of  $y_0$  will gradually increase, and the volume of  $V_{z1}$  will decrease at this time, indicating that the proportion of e-commerce platform to implement information supervision decreases.  $\square$

**3.4.4. Strategy Stability Analysis of Government Regulatory Department.** Assuming that the expected benefit of government regulatory department when government implementing the strategy of strictly supervising is  $U_{41}$ , the expected benefit of government regulatory department when government implementing the strategy of loosely supervising is  $U_{42}$ , and the average expected benefit of the government regulatory department is  $\overline{U}_4$ , which are defined as follows:

$$\begin{aligned} U_{41} &= xyz(S - C_g + R) + x(1-y)z(S - C_g + R) \\ &\quad + (1-x)yz(S - C_g + I_e) + (1-x)(1-y)z(S - C_g + I_e) \\ &\quad + xy(1-z)(S - C_g + R) + x(1-y)(1-z)(S - C_g + R) \\ &\quad + (1-x)y(1-z)(S - C_g + I_e + I_p) \\ &\quad + (1-x)(1-y)(1-z)(S - C_g + I_e + I_p) \\ &= S - C_g + xR + (1-x)I_e + (1-x)(1-z)I_p. \\ U_{42} &= xyzS + x(1-y)zS + (1-x)yz(S - N + \alpha I_e) \\ &\quad + (1-x)(1-y)z(S - N) + xy(1-z)S \\ &\quad + x(1-y)(1-z)S + (1-x)y(1-z)(S - N + \alpha I_e + \alpha I_p) \\ &\quad + (1-x)(1-y)(1-z)(S - N) \\ &= S - (1-x)N + (1-x)y\alpha[I_e + (1-z)I_p], \\ \overline{U}_4 &= rU_{41} + (1-r)U_{42}. \end{aligned} \quad (13)$$

According to the Malthusian dynamic equation, the replication dynamic equation of the government regulatory department is obtained as follows:

$$\begin{aligned} F(r) &= \frac{dr}{dt} \\ &= r(U_{41} - \overline{U}_4) \\ &= r(1-r)\{-C_g + xR + (1-x)I_e + (1-x)(1-z)I_p \\ &\quad + (1-x)N - (1-x)y\alpha[I_e + (1-z)I_p]\}. \end{aligned} \quad (14)$$

The first partial derivative of  $F(r)$  for  $r$  is as follows:

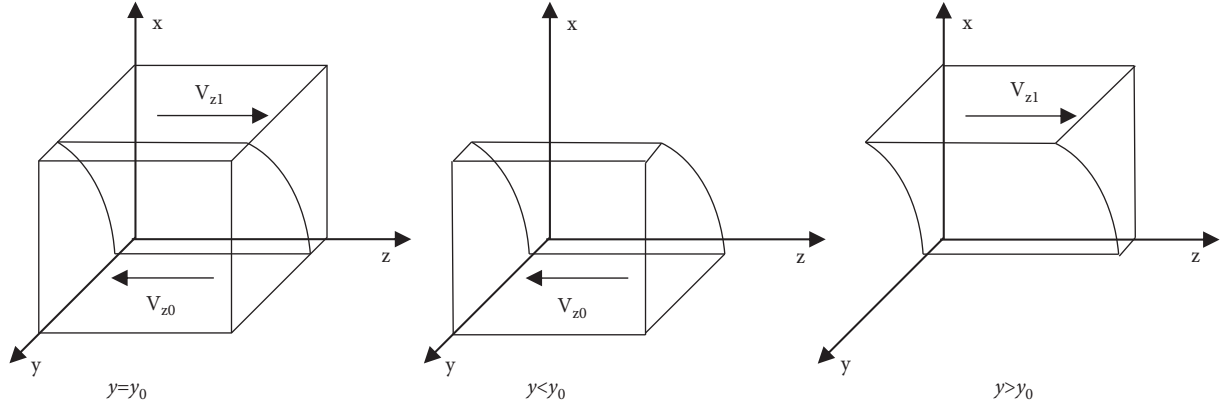


FIGURE 4: Phase diagram of strategy evolution of e-commerce platform.

$$F'_r(r) = (1 - 2r) \left\{ -C_g + xR + (1 - x)I_e + (1 - x)(1 - z)I_p + (1 - x)N - (1 - x)y\alpha [I_e + (1 - z)I_p] \right\}. \quad (15)$$

Based on the stability theorem of differential equations, government regulatory department implements the strategy of strictly supervising in the stable state must meet the conditions:  $F(r) = 0$ , and  $F'_r(r) < 0$ .

**Proposition 4.** When  $z > z_0$ , the government regulatory department will choose strict supervision as the stable strategy; when  $z < z_0$ , the stable strategy of the government regulatory department will choose loose supervision as the stable strategy; when  $z = z_0$ , the government regulatory department cannot determine the stable strategy. Where the threshold is as follows:

$$z_0 = \frac{-C_g + xR + (1 - x)(1 - \alpha y)I_p + (1 - x)[(1 - \alpha y)I_e + N]}{(1 - x)(1 - \alpha y)I_p}. \quad (16)$$

*Proof.* Assume  $H(z) = -C_g + xR + (1 - x)I_e + (1 - x)(1 - z)I_p + (1 - x)N - (1 - x)y\alpha [I_e + (1 - z)I_p]$ , when  $(\partial H/\partial x) < 0$ ,  $H(z)$  is considered to be an increasing function of  $z$ . When  $z < z_0$ ,  $H(z) > 0$ ,  $F(r)|_{r=1} = 0$ , and  $F'_r(r)|_{r=1} < 0$ , so  $r = 1$  has stability; When  $z > z_0$ ,  $H(z) < 0$ ,  $F(r)|_{r=0} = 0$ , and  $F'_r(r)|_{r=0} < 0$ , so  $r = 0$  has stability; When  $z = z_0$ ,  $H(z) = 0$ ,  $F(r) = 0$ , and  $F'_r(r) = 0$ , so  $z$  is stable at all levels in the range of 0 to 1, that is, the government regulatory department's strategy does not change over time, regardless of the proportion of government regulatory department choosing to strict supervision.

Proposition 4 states that the decline of the proportion of information supervision of e-commerce company by e-commerce platform will change the stable strategy of government regulatory department from loose supervision to strict supervision; Similarly, the increase of the proportion of information supervision of e-commerce company by e-commerce platform will change the stable strategy of government regulatory department from strictly supervising to loosely supervising. Therefore, the government's strict

supervision on e-commerce company is the necessary measure under the unfavorable conditions of the e-commerce platform's information supervision on e-commerce company.

Based on Proposition 4, the phase diagram of strategy evolution of the government regulatory department is shown in Figure 5.

*Inference 4:* With the increase of the value of  $R$ ,  $I_e$ ,  $I_p$ , and  $N$ , the government regulatory department is more inclined to implement the strict supervision strategy, when other parameters remain unchanged. Similarly, with the increase of the value of  $C_g$  and  $\alpha$ , the government is more inclined to implement the loose supervision strategy. It shows that the proportion of government regulatory department implementing strict supervision strategy is directly proportional to the social benefits obtained, the fines punished by the government on e-commerce company and platform, and the social reputation loss caused by differential pricing under the government's loose supervision, and inversely proportional to the cost for the government to strictly supervise and the proportion of consumer discovering differential pricing.  $\square$

*Proof.* Since  $z_0 = 1 - (C_g - xR - (1 - x)[(1 - \alpha y)I_e + N]) / ((1 - x)(1 - \alpha y)I_p)$ , the volume of  $V_{r1}$  in Figure 5 represents the proportion of strictly supervised by the government, and the corresponding volume of  $V_{r0}$  represents the proportion of loosely supervised by government. When the value of  $R$ ,  $I_e$ ,  $I_p$ , and  $N$  gradually increases, the value of  $z_0$  will gradually increase, and the volume of  $V_{r1}$  will increase at this time, indicating that the proportion of strict supervision by government regulatory department increases; When the value of  $C_g$  and  $\alpha$  gradually increase, the value of  $z_0$  will gradually decrease, and the volume of  $V_{r1}$  will decrease at this time, indicating that the proportion of strict supervision by government regulatory department increases decreases.  $\square$

## 4. Results and Discussion

*4.1. ESS Analysis among Four-Party Game Players.* In the dynamic system of government regulatory department,

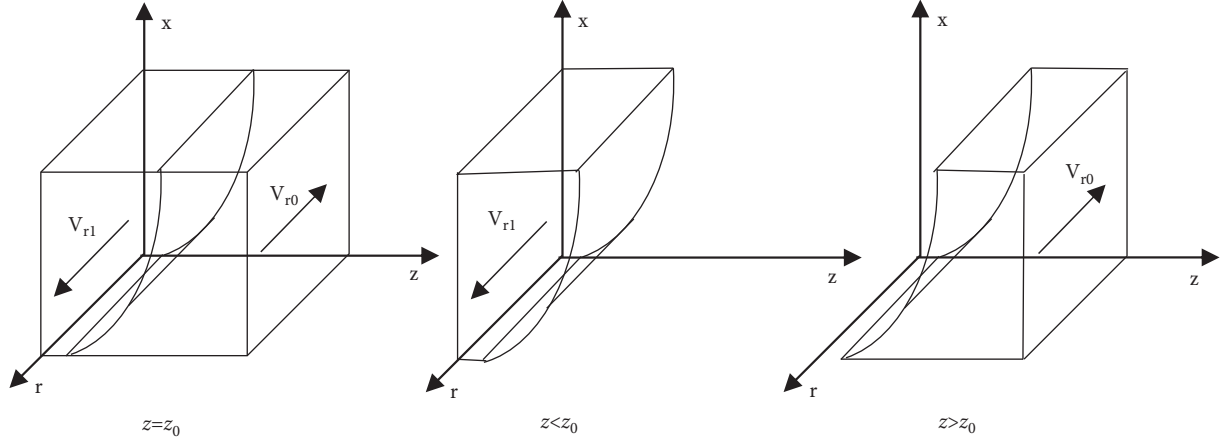


FIGURE 5: Phase diagram of strategy evolution of government regulatory department.

e-commerce platform, e-commerce company and consumer, the stability of the strategic combination of the four-party game subjects can be referred to as the nonlinear function stability discriminant method of First Law of Lyapunov. Ritzberger and Weibull [38] and Selten [39] pointed out that the stable solutions in the multi-group evolutionary game are strict Nash equilibrium, which must be the pure strategy. Therefore, this study analyzes 16 pure strategies in four-party evolutionary game learning from the research method of Sun and Su [40].

Due to the replication dynamic equation of each game subject, the Jacobian matrix is obtained as follows:

$$J = \begin{bmatrix} F'_x(x) & F'_y(x) & F'_z(x) & F'_r(x) \\ F'_x(y) & F'_y(y) & F'_z(y) & F'_r(y) \\ F'_x(z) & F'_y(z) & F'_z(z) & F'_r(z) \\ F'_x(r) & F'_y(r) & F'_z(r) & F'_r(r) \end{bmatrix}, \quad (17)$$

where the elements in the matrix are shown in Appendix A.

#### 4.1.1. ESS Analysis among Four-Party Game Players under the Strict Supervision of Government Regulatory Department.

When  $C_g - xR - (1-x)I_e - (1-x)(1-z)I_p - (1-x)N + (1-x)\gamma\alpha[I_e + (1-z)I_p] < 0$ , government regulatory department implements strict supervision. According to the Jacobian matrix shown in Appendix B, the equilibrium solution of the four-party evolutionary game can be obtained, and the stability analysis is shown in Table 3.

Condition (a):  $-P_n + P_d + \Delta P - M - I_e < 0$ ,  $C_g - R < 0$ , and  $-C_p + T_p < 0$

Condition (b):  $-P_n + P_d + \Delta P - M - I_e - F < 0$ ,  $C_g - R < 0$ , and  $C_p - T_p < 0$

It can be seen from Table 3 that there are two possible stable strategies under strict supervision by the government regulatory department, i.e. E5 (1, 1, 0, 1) and E8 (1, 1, 1, 1).

When the condition (a) is met, that is,  $-P_n + P_d + \Delta P - M - I_e < 0$ ,  $C_g - R < 0$ , and  $-C_p + T_p < 0$ . The sum of the benefits of differential pricing to loyal consumers by

TABLE 3: Asymptotic stability analysis of equilibrium point of replication dynamic system under the strict supervision of government regulatory department.

Equilibrium point	Eigenvalue symbol	Stability of equilibrium point
E1 (0, 0, 0, 1)	(+, X, X, X)	Instability point
E2 (1, 0, 0, 1)	(-, +, -, -)	Instability point
E3 (0, 1, 0, 1)	(+, X, X, -)	Instability point
E4 (0, 0, 1, 1)	(+, X, X, -)	Instability point
E5 (1, 1, 0, 1)	(-, -, -, -)	ESS in condition (a)
E6 (1, 0, 1, 1)	(-, +, +, -)	Instability point
E7 (0, 1, 1, 1)	(+, X, X, -)	Instability point
E8 (1, 1, 1, 1)	(-, -, -, -)	ESS in condition (b)

Note: X means uncertain of symbol, and ESS means the evolutionarily stable strategy.

e-commerce company is less than the sum of the benefits of nondifferential pricing by e-commerce company to the consumer and the fines to the e-commerce company for differential pricing and compensation of e-commerce company to consumer by the government. The strict supervision cost is less than the social benefits when controlling differential pricing for the government. And the reputation value produced by the loyal consumer to the platform is less than the cost of the platform information supervision. Then the strategy of each subject is stable at equilibrium point E5 (1, 1, 0, 1). E-commerce company implements nondifferential pricing, the consumer is loyal to the e-commerce company, e-commerce platform implements information nonsupervision, and the government strictly supervises e-commerce platform and e-commerce company. This situation may exist in the period of chaotic pricing for the e-commerce company. Since the e-commerce platform benefits less from the information supervision of e-commerce company, it has no motivation to supervise e-commerce company. Therefore, the government must come forward to supervise differential pricing, safeguard consumer rights and interests, and help e-commerce company gain consumer loyalty.

When the condition (b) is met, that is  $-P_n + P_d + \Delta P - M - I_e - F < 0$ ,  $C_g - R < 0$ , and  $C_p - T_p < 0$ . With the

improvement of consumers' awareness of differential pricing and the reduction of the cost of platform supervising information, the information supervision cost of the platform is less than the reputation value brought by the loyal consumer to the platform, and the e-commerce platform can also join into the supervision of e-commerce company. When the other conditions remain unchanged, the strategy of each subject is stable at equilibrium point E8 (1, 1, 1). The government and e-commerce platform jointly strengthen the supervision of differential pricing of e-commerce company, so that e-commerce company inclined to to be nondifferential pricing, and consumer is loyal to the e-commerce company.

#### 4.1.2. ESS Analysis among Four-Party Game Players under the Loose Supervision of Government Regulatory Department.

When  $C_g - xR - (1-x)I_e - (1-x)(1-z)I_p - (1-x)N + (1-x)y\alpha[I_e + (1-z)I_p] > 0$ , government regulatory department implements loosely supervision. According to the Jacobian matrix shown in Appendix C, the equilibrium solution of the four-party evolutionary game can be obtained, and the stability analysis is shown in Table 4.

Condition (c):  $-P_n + P_d + \Delta P - \alpha(M + I_e) - F < 0$ ,  $C_p - T_p < 0$ , and  $-C_g + R < 0$ .

As shown in Table 4 that there is a possible stabilization strategy under loose supervision by government regulatory authorities, i.e. E16 (1, 1, 1, 0).

When the condition (c) is met, that is,  $-P_n + P_d + \Delta P - \alpha(M + I_e) - F < 0$ ,  $C_p - T_p < 0$ , and  $-C_g + R < 0$ . The sum of the benefits of e-commerce company's differential pricing for the loyal consumer is less than the sum of the benefits of e-commerce company's nondifferential pricing for consumer, the fines punished by government regulatory department under loosely supervising and the compensation for the consumer for differential pricing of e-commerce company, and the fines imposed by e-commerce platform on the e-commerce company. The reputation value brought by the loyal consumer to the platform is greater than the cost of the platform information supervision. And the strict supervision cost is greater than the social benefits when controlling differential pricing for the government. Then the strategy of each subject is stable at equilibrium point E16 (1, 1, 1, 0). This situation may exist in the normative period of discriminatory pricing by the e-commerce company. At this time, as the proportion of the differential pricing of e-commerce company gradually decreases, the social benefits of the government's strict supervision of differential pricing decrease. When the social benefit is less than the strictly supervising cost of the government regulatory department, the strategy of the government regulatory department will change from strictly supervising to loosely supervising. The main responsibility of supervision will be transferred from the government to the e-commerce platform and consumer. Supervision and fines by e-commerce platform and consumer enable e-commerce company to conduct nondifferential pricing and promote the virtuous circle of the e-commerce industry ecosystem.

TABLE 4: Asymptotic stability analysis of equilibrium point of replication dynamic system under the loose supervision of government regulatory department.

Equilibrium point	Eigenvalue symbol	Stability of equilibrium point
E9 (0, 0, 0, 0)	(+, X, X, X)	Instability point
E10 (1, 0, 0, 0)	(-, +, X, X)	Instability point
E11 (0, 1, 0, 0)	(+, X, X, X)	Instability point
E12 (0, 0, 1, 0)	(+, X, X, X)	Instability point
E13 (1, 1, 0, 0)	(X, -, X, +V)	Instability point
E14 (1, 0, 1, 0)	(X, +, +, +)	Instability point
E15 (0, 1, 1, 0)	(+, X, X, X)	Instability point
E16 (1, 1, 1, 0)	(-, -, -, -)	ESS in condition (c)

Note: X means uncertain of symbol, and ESS means the evolutionarily stable strategy.

4.2. Numerical Simulation Analysis. In order to test the reliability of the model and more intuitively demonstrate the influence of key factors in the replication dynamic system on the evolutionary trajectory of stakeholders of the multi-party game, the model is given numerical value combined with the actual situation, and the numerical simulation is carried out by MATLAB2021.

For the e-commerce company operating in the e-commerce platform, the benefit of nondifferential pricing to the consumer is set as  $P_n = 10$ , and the benefit of differential pricing to the consumer is set as  $P_d = 9$ , and the additional benefit of differential pricing to the loyal consumer is set as  $\Delta P = 5$ . If differential pricing is discovered by the government, the compensation of the e-commerce company to the consumer is set as  $M = 4$ . The reputation value brought by the loyal consumer to the e-commerce company is set as  $T_e = 5$ , and the reputation value brought by the loyal consumer to the e-commerce platform is set as  $T_p = 5$ . The utility obtained by the loyal consumer when purchasing goods from the e-commerce company is set as  $U_l = 12$ , and the utility obtained by the disloyal consumer when purchasing goods from the e-commerce company is set as  $U_d = 11$ . The probability of loyal consumer discovering differential pricing under government loose supervision is set as  $\alpha = 0.2$  and the complaint cost of the loyal consumer is set as  $C_c = 3$ . The social benefit of nondifferential pricing obtained by the government under strict supervision is set as  $R = 7$ , and the cost of strictly supervised by the government is set as  $C_g = 6$ . The fine by government regulatory department for differential pricing of e-commerce company  $I_e = 3$ . The social reputation loss of the government caused by differential pricing under loose supervision is set as  $N = 8$ . The normal benefit obtained by the government from the operation of the platform is set as  $S = 6$ . The benefit of the platform reasonably providing information to e-commerce company is set as  $W = 5$ , and the cost of the platform's information supervision on e-commerce company is set as  $C_p = 7$ . The fine imposed by the platform to e-commerce company for differential pricing during information supervision is set as  $F = 3$ , and the proportion of fine imposed by the e-commerce platform for differential pricing of the e-commerce company is set as  $\beta = 0.6$ .



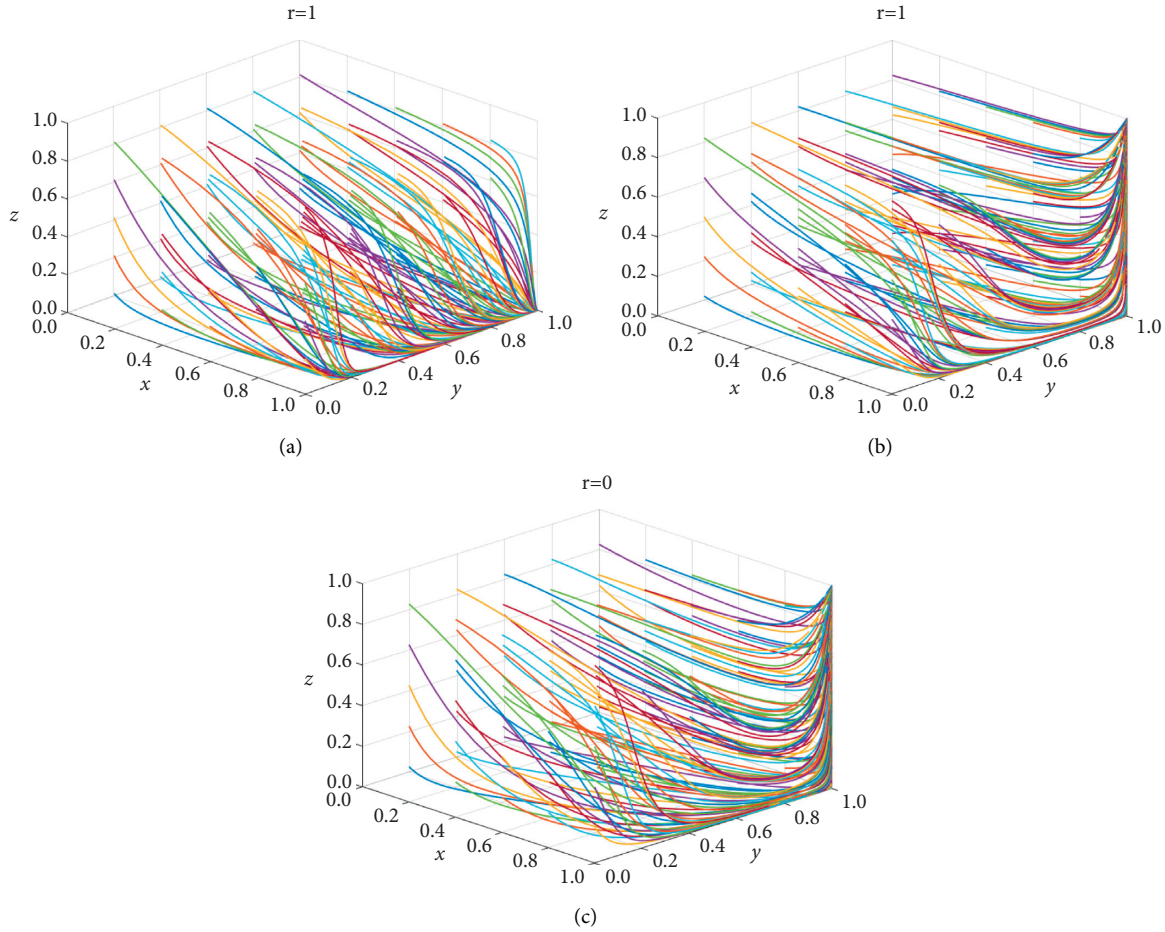


FIGURE 6: Influence of the establishment of government supervision mechanism on strategy evolution of all parties.

**4.2.1. The Influence of Government Supervision Mechanism.** To test whether the government supervision mechanism is effective in the process of differential pricing of e-commerce company, the proportions of government strict supervision are set as  $r=0$  and  $r=1$  to represent the two states of loose supervision and strict supervision of government supervision department. The evolution process of different initial strategies of the e-commercial company, consumer, and e-commerce platform is simulated and analyzed in three-dimensional space, and the simulation results with time are shown in Figure 6.

As shown in Figure 6(a) that when government regulatory department adopts the strict supervision strategy on the differential pricing of e-commerce company, although the e-commerce platform does not take information supervision strategy on account of the high cost for information supervision, the strategies of the e-commerce company and consumer can still incline to be stable in nondifferential pricing and loyalty. This shows that it is very necessary and effective for the government to adopt the strict supervision strategy. With the reduction of  $C_p$ , that is, the information supervision cost reduced, the platform will be inclined to adopt the strategy of information supervision, to achieve coordinated supervision to e-commerce company by the government and platform, then the company adopts

nondifferential pricing, and consumer is loyal to the e-commerce company. And the stable strategy portfolio is demonstrated in Figure 6(b). As is exhibited in Figure 6(c) that when government regulatory department implements the loosely supervising to e-commerce company for the differential pricing due to the high cost of strict supervision, if  $C_p$  is small, that is, the cost of information supervision on the e-commerce platform is small, and  $\alpha$  is at a high level, the consumer can actively discover the differential pricing of the e-commerce company and report it, the e-commerce company will also incline to nondifferential pricing. Therefore, although the government selects the loose supervision strategy, the differential pricing behavior of e-commerce company is supervised collaboratively by the platform and consumer. The strategy equilibrium is consistent with the previous analysis of the stability under different government supervision strategies.

**4.2.2. The Influence of Information Supervision Cost of E-Commerce Platforms.** If  $C_p = \{7, 4, 1\}$ , the stability of the system evolution of the four-party game subjects and the simulation results are shown in Figure 7.

According to Figure 7, with the reduction of the information supervision cost of the e-commerce platform, the



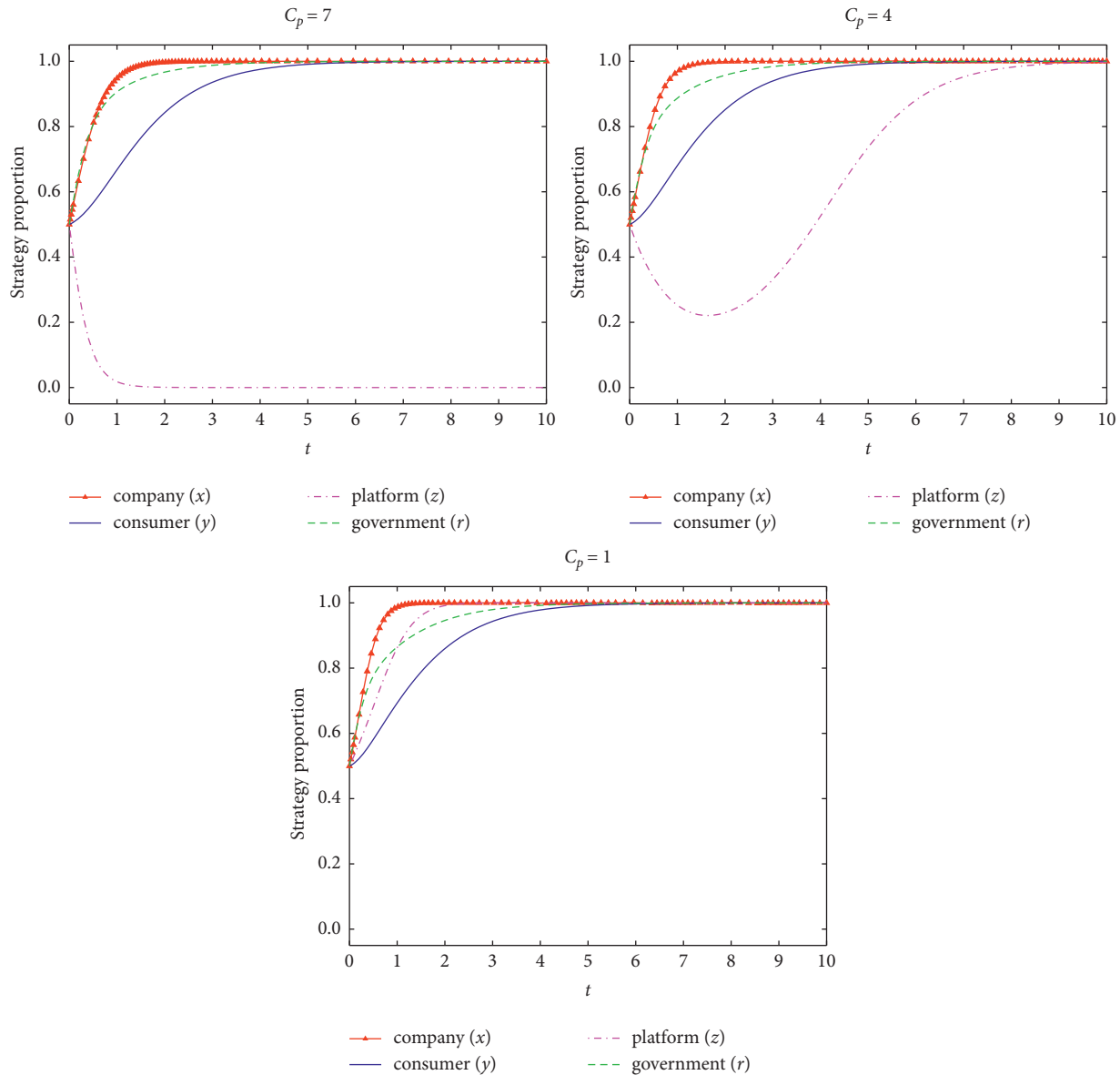


FIGURE 7: Influence of information supervision cost of e-commerce platform on strategy evolution of all parties.

supervision strategy of the platform will be transformed from information nonsupervision on e-commerce company to information supervision. Therefore, the platform can join the ranks of the government to regulate the company, and collaboratively supervise the differential pricing of the e-commerce company for loyal consumer. Moreover, the less the information supervision cost of the platform, the faster the stable strategy of information supervision. Therefore, active measures can be adopted to lower the cost for information supervising of e-commerce platform, to stimulate e-commerce platform to supervise the differential pricing behavior of e-commerce company on the platform.

4.2.3. *The Influence of the Strict Supervision Cost of Government Regulatory Department.* If  $C_g = \{6, 8, 10\}$ , the stability of the system evolution of the four-party

game subjects and the simulation results are shown in Figure 8.

According to Figure 8, the strict supervision cost of government affects the decision-making of government regulatory department, as well as affects the evolution of decision-making of the other subjects. With the increase of government supervision cost, the supervision strategy of the government regulatory department to the differential pricing of e-commerce company will be transformed from strict supervision to loose supervision, and gradually become the cyclical alternating strategy between strict supervision and loose supervision with medium proportion. The strategy of the e-commerce platform will be also transformed from information supervision to information nonsupervision of e-commerce company when strictly supervising cost of government Increasing. Free from the supervision of government regulatory department and platform, the pricing strategy of the company for the loyal

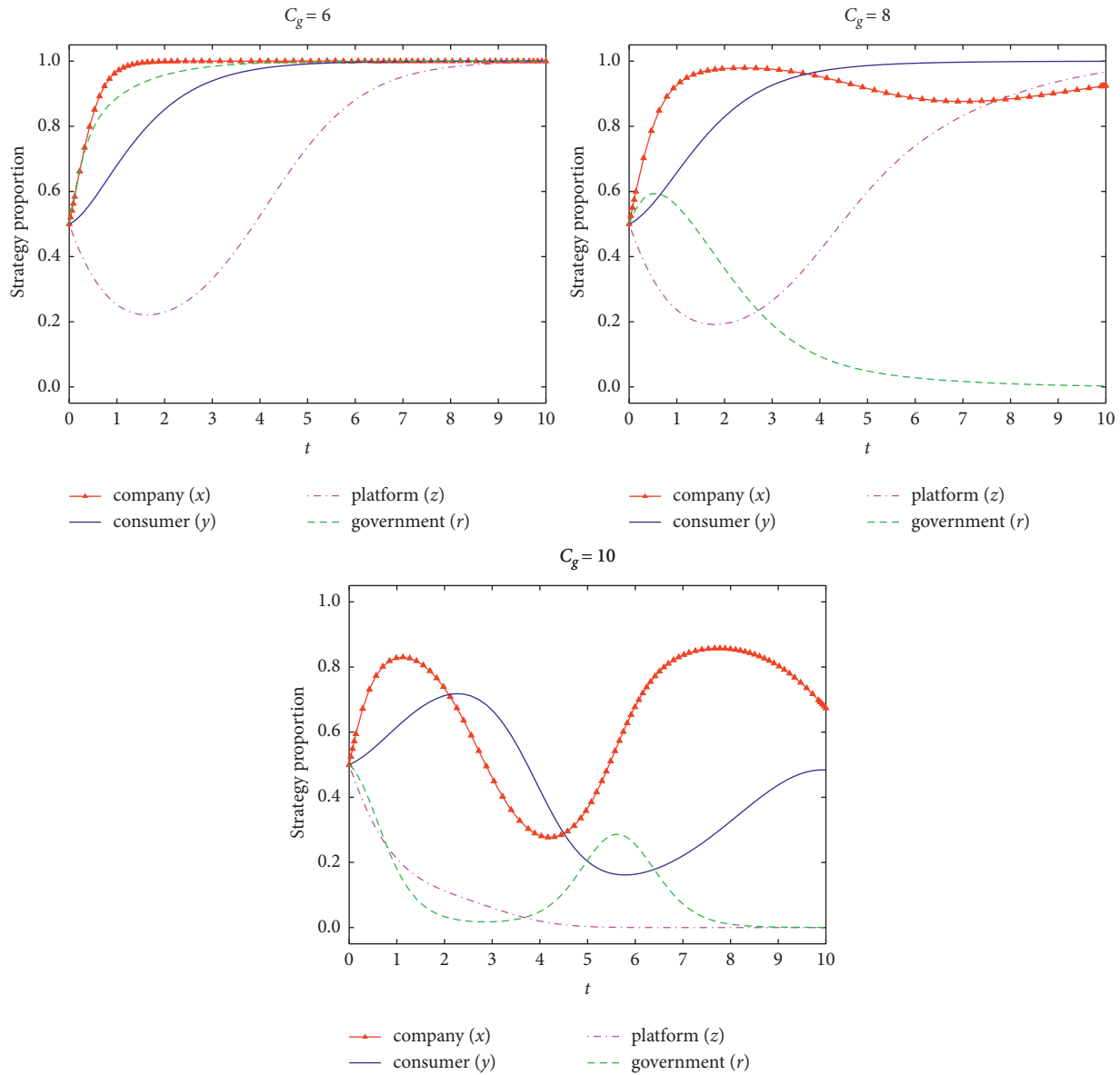


FIGURE 8: Influence of strict supervision cost of government regulatory department on strategy evolution of all parties.

consumer will be transformed from nondifferential pricing to moderate-proportion differential pricing, and the strategy change periodically. With the increase of the strictly supervising cost of government, the strategy of the consumer will be transformed from loyalty to e-commerce company to disloyalty. Therefore, the strict supervision cost of the government regulatory department is the key factor in restricting the differential pricing of the e-commerce company. Measures should be arranged to actively reduce the strictly supervising cost of the government regulatory department at a certain level, to stimulate platform and the consumer to regulate the behavior of e-commerce company in differential pricing.

4.2.4. The Influence of the Probability of Loyal Consumer Discovering Differential Pricing under Government’s Loose Supervision. If  $\alpha = \{0.1, 0.3, 0.5\}$ , the stability of the system

evolution of the four-party game subjects and the simulation results are shown in Figure 9.

According to Figure 9, with the increase of probability of loyal consumer discovering differential pricing under government loose supervision, the probability of exposure of differential pricing behavior of e-commerce company for loyal consumer increases, which will make e-commerce company gradually improve the proportion of nondifferential pricing and stabilize in the nondifferential pricing strategy. The e-commerce platform can also gradually improve the proportion of information supervision due to the increase of fines for nonsupervision of e-commerce company information resulting in differential pricing, and the behavior stabilizes in the information supervision strategy. The government regulatory department can gradually loose supervision and transfer the responsibility of supervision to e-commerce platform and

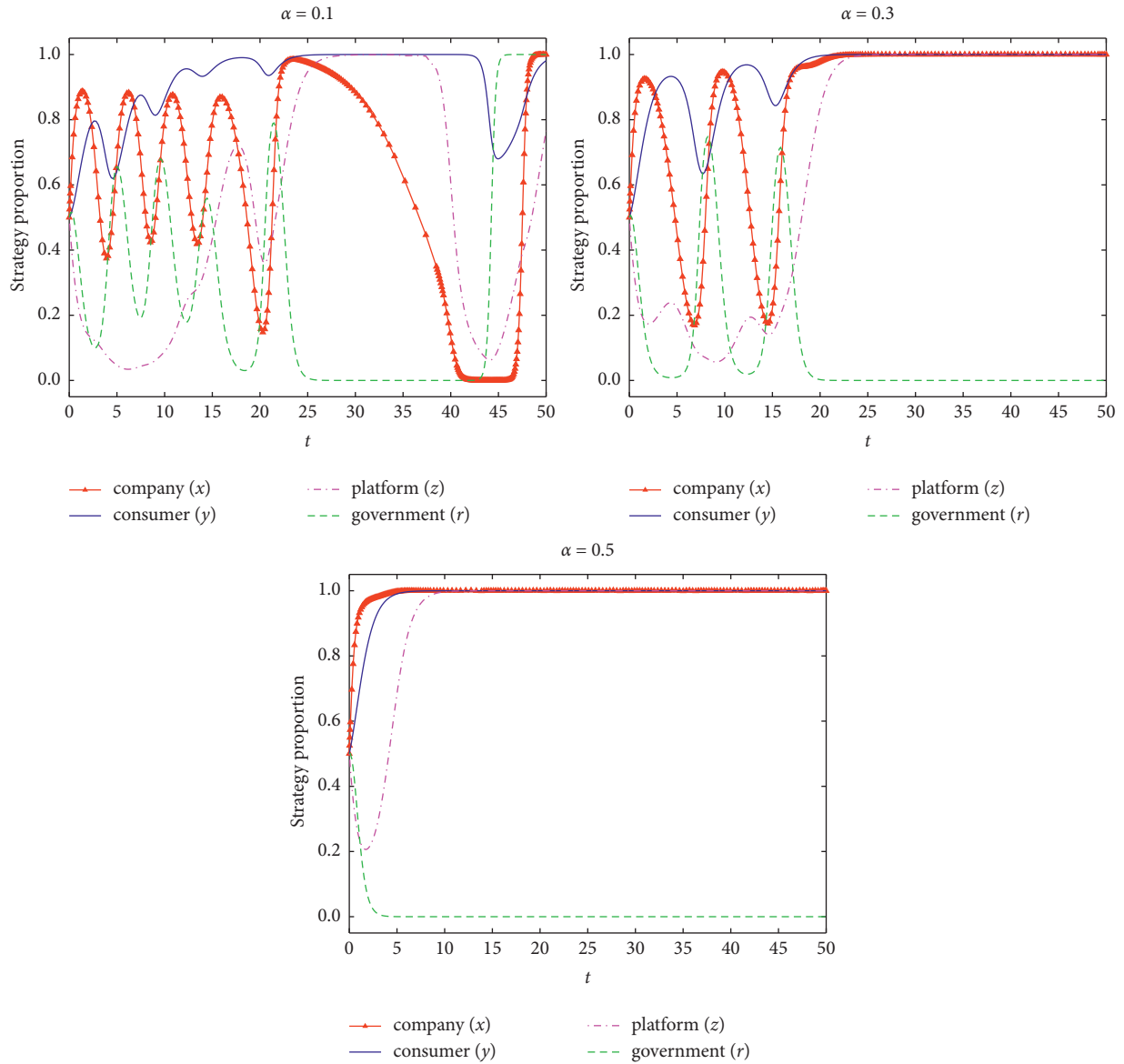


FIGURE 9: Influence of the probability of loyal consumer discovering differential pricing under government loose supervision on strategy evolution of all parties.

the consumer. Therefore, measures can be taken to encourage the consumer to report the differential pricing behavior of e-commerce company, to maintain the stable and sustainable progress of e-commerce platform and systems.

**4.2.5. The Influence of the Penalties for Differential Pricing of E-Commerce Company under Government's Loose Supervision.** If  $M = \{1, 2, 4\}$ ,  $I_e = \{1, 2, 4\}$ , and  $F = \{1, 2, 4\}$ , the evolution process and results of the strategy of the four-party game subjects are shown in Figure 10.

According to Figure 10, with the increase of the fines given by consumer, e-commerce platform, and government regulatory department for differential pricing of e-commerce company, the e-commerce company will gradually

increase the proportion of nondifferential pricing and stabilize in the nondifferential pricing strategy. The consumer will increase the proportion of loyalty to the e-commerce company and the behavior stabilize in the loyalty strategy when the compensation for differential pricing from e-commerce company increases to compensate for the loss of differential pricing. The e-commerce platform will also gradually improve the proportion of information supervision due to the increase of benefits from information supervision fines and the behavior stabilizes in the information supervision strategy. Therefore, the nondifferential pricing behavior of e-commerce company can be promoted by increasing the punishment for differential pricing, to realize the joint dynamic supervision of the e-commerce platform, the consumer, and the government on the pricing of the e-commerce company.

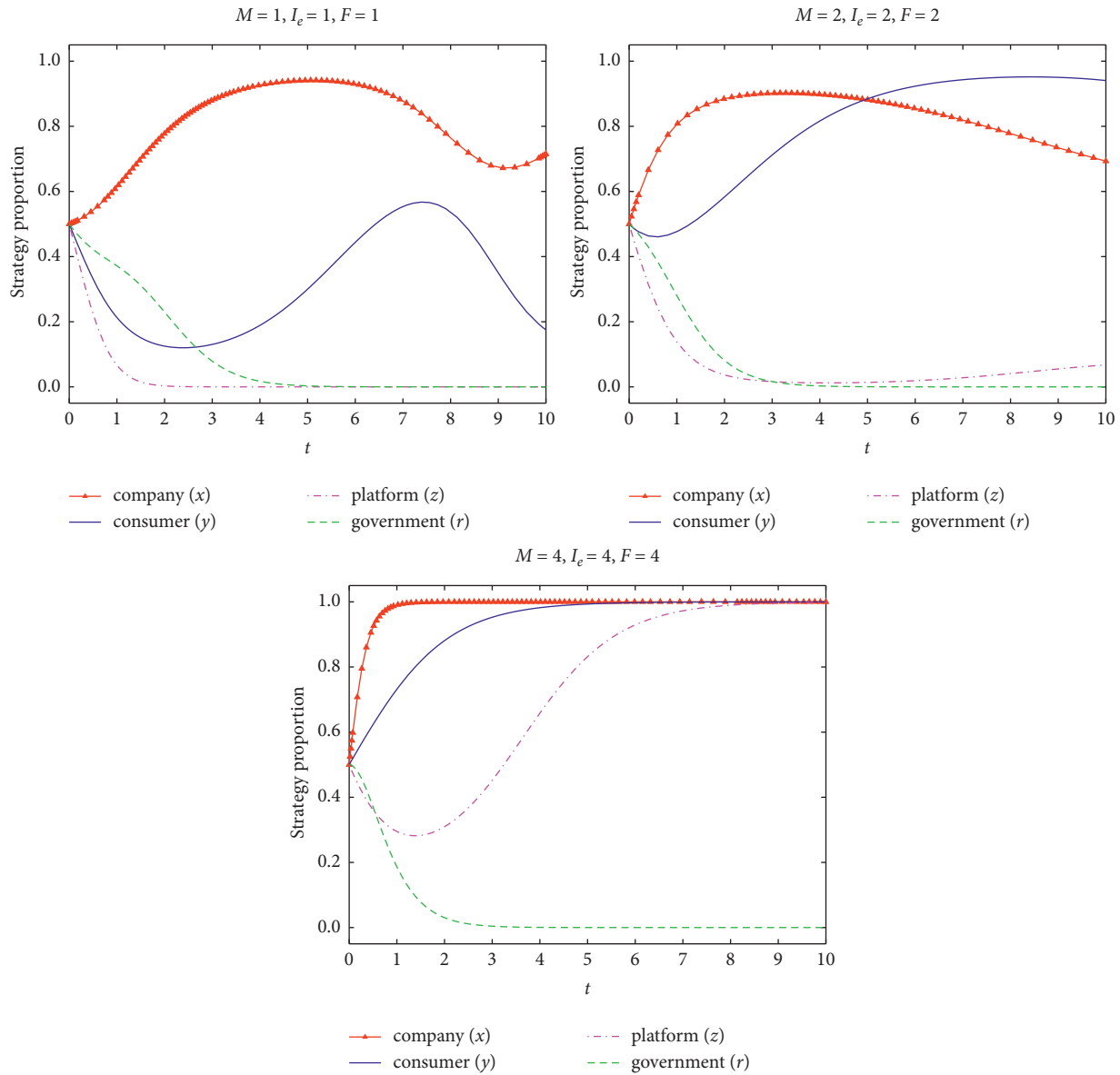


FIGURE 10: Influence of the penalties for differential pricing of e-commerce company under government loose supervision on strategy evolution of all parties.

### 5. Conclusions

Given the phenomenon of “big data killing” that e-commerce companies use customer information in the pricing process, this paper studies how to safeguard consumers’ pricing fairness in the context of the Internet, and builds the four-party evolutionary game model for the supervision on differential pricing of e-commerce company, analyzes the stability of the strategy selection of each subject in the model, and the stability of equilibrium point of the strategic combination in the replication dynamic system, and simulates and analyzes the influence of key elements on the strategy evolution. The main conclusions are as follows:

- (1) The government supervision mechanism can play an effective role to limit differential pricing of the e-commerce company. When the proportion of

strict government supervision adds, the sum of the benefits of differential pricing for loyal consumers by e-commerce company is less than the penalty cost of e-commerce company, and strict supervision cost of government is less than its social benefits, then e-commerce company inclines more to choose the strategy of nondifferential pricing. Since the reputation value of the e-commerce platform is less than the information supervision cost of platform, the platform inclines more to conduct information nonsupervision. Therefore, the equilibrium strategy of each subject is stable at point E5 (1, 1, 0, 1), which occurs in the early stage of the government’s strict supervision on the e-commerce company. With the reduction of the supervision cost of the platform, it is also willing to join the supervision on differential

pricing of e-commerce company for the platform and inclines more to choose information supervision strategy. Therefore, the equilibrium strategy of each subject is stable at point E8 (1, 1, 1, 1), which occurs in the stable stage of the government's strict supervision of e-commerce company, and the participation of the e-commerce platform relieved the pressure on government supervising on the company. When the strictly supervising cost of government increases, the reputation value of the platform is greater than the supervision cost of the platform, then the government regulatory department inclines more to loose supervision strategy. Therefore, the equilibrium strategy of each subject is stable at point E16 (1, 1, 1, 0), which occurs in the later stage of the government's strict supervision of e-commerce company. When both e-commerce platform and consumer realize the important role of supervision and conduct strong collaborative supervision, the government can take the way of auxiliary supervision to control the differential pricing of the e-commerce company.

- (2) The information supervision cost of the e-commerce platform is the main factor affecting the supervision strategy of the platform. When the supervision cost of the platform is greater than the reputation value of the platform, the platform inclines more to conduct information nonsupervision. However, as the information supervision cost of the platform decreases and is less than the reputation value of the platform, the stable strategy of the e-commerce platform will transform into information supervision and then promote nondifferential pricing for the e-commerce company. Moreover, the less the information supervision cost of the e-commerce platform, the faster the stable strategy of the e-commerce platform can transform into the information supervision strategy.
- (3) The strict supervision cost of government is the main factor affecting the strategies of all parties. When the strict supervision cost of government is so small as to be less than the social benefits of government strict supervision on differential pricing of e-commerce company, the equilibrium strategy of all parties is that both government regulatory department and e-commerce platform implement supervision, e-commerce company conduct nondifferential pricing, and consumer is loyal to the e-commerce company. However, when the strict supervision cost of government increases and exceeds the social benefits of government strict supervision on differential pricing of e-commerce company, the government gradually inclines to loose supervision strategy. At this moment, if platform and consumer can supervise the pricing of the e-commerce company to a certain extent, e-commerce company still incline to nondifferential pricing strategy. When the strict supervision cost of government increases to a very high level, not only the government cannot

strictly supervise, but also e-commerce platform will not supervise the information used by the e-commerce company. Then e-commerce company will incline to differentiate pricing, and the consumer will be disloyal.

- (4) The probability of the consumer discovering differential pricing under the government's loose supervision policy is an important factor affecting the strategies of all parties. As strict supervision cost of government is at a higher level, and the probability of consumer discovering differential pricing of the e-commerce company is small, neither the government nor the e-commerce platform can incline to the more stable behavioral strategy. Although customer inclines to be loyal to the e-commerce company, the strategies of four subjects cannot maintain the stable equilibrium, and the strategy of e-commerce company become cyclical alternating between differential pricing and nondifferential pricing. When the probability of consumer discovering differential pricing of e-commerce company increases, e-commerce company gradually inclines to nondifferential pricing strategy, e-commerce platform gradually inclines to information supervision strategy, and the government gradually inclines to loose supervision strategy. The equilibrium strategy of four-party behavior achieves. The higher the probability level of consumer discovering differential pricing, the faster the equilibrium strategy of four-party behavior achieves. This conclusion also confirms the conclusion in the research of Yu and Li [9] and Wu et al. [30] that the probability of consumer finding himself killed in price is the important factor affecting the strategy choice of consumer and company.
- (5) The penalties for differential pricing of e-commerce company under the government's loose regulatory are the important factors affecting the strategies of all parties. When consumers, e-commerce platform and government regulatory department impose the fines and compensation on differential pricing of e-commerce companies at a low level, the government and e-commerce platform incline to not supervise, and e-commerce company and consumer cannot maintain a stable equilibrium. When the penalties for differential pricing of e-commerce company is high, e-commerce platform inclines to supervise the information, consumer inclines to be loyal, while e-commerce company inclines to price nondifferentially, the government inclines to loose supervise, and the strategies remains stable. The higher the penalties for differential pricing of e-commerce company, the faster the equilibrium strategy of four-party behavior achieves.

In this study, the modeling analysis and simulation of the supervision of "big data killing" of e-commerce company are carried out, which breaks through the limitation of analyzing only two or three parties in the existing "big data killing"

problem. It is a beneficial supplement to systematic research on this issue that more participants consider their action strategies under the same system. The four-party evolutionary game model constructed also expands the application scope of the evolutionary game method in the study of pricing supervision of e-commerce company. The research conclusions can provide favorable theoretical support for the “big data killing” problem in practice.

Therefore, to better restrain the pricing behavior of e-commerce company, regulate differential pricing, and build a good e-commerce shopping environment, the following measures should be taken by the government regulatory department, e-commerce platform, e-commerce company, and consumer.

- (1) From the perspective of the government, the government regulatory department must supervise e-commerce company, especially in the early stage of price discrimination by using customer information. Therefore, the government needs to use economic and policy means to effectively manage the operation of e-commerce platform and company and promote the enthusiasm of e-commerce company to conduct nondifferential pricing. For example, adopting more advanced big data analysis technology to supervise price changes of e-commerce company; establishing more extensive and efficient reporting channels so that consumers can timely price complaints; improving corresponding legal measures to increase the violation cost of the e-commerce company and punishing “big data killing” from the aspects of economy and reputation. While supervising, it is also necessary to pay attention to reducing the strict supervision cost of the government regulatory department.
- (2) From the perspective of the e-commerce platform, as the important carrier of e-commerce operation, the e-commerce platform should strengthen information supervising of the e-commerce company. The e-commerce platform is the main body that controls customer information. E-commerce company conducts “big data killing” differential pricing based on the mastery of customer information. Therefore, the e-commerce platform needs to carry out information supervision when providing information for the e-commerce company and formulates policies to punish e-commerce company with differential pricing. It is also necessary to improve the technical management level of the e-commerce platform, and use innovative technology based on big data to monitor e-commerce company and reduce the supervision cost of the e-commerce platform.
- (3) From the perspective of consumers, they should actively protect their rights and interests. While online shopping brings convenience to consumers, it may also lead to the possibility of price discrimination with consumer information. In the process of e-commerce shopping, consumers will prefer some

e-commerce companies due to path dependence, and then form customer loyalty, but this path dependence should not be the reason for the differential pricing of e-commerce companies. Therefore, consumers should enhance price sensitivity and verify the displayed price of e-commerce companies through various channels, to reduce the infringement of consumer rights and interests by e-commerce companies.

- (4) From the perspective of the e-commerce company, although maximizing profits is the important motive of business behavior, the reputation and service in e-commerce shopping are the foundation for the long-term development of the e-commerce company. Under the market conditions where consumers’ transfer costs are getting lower and lower, the e-commerce company can grow gradually mainly based on gaining the loyal customer. Therefore, e-commerce company should not adopt the differential pricing strategy in pursuit of temporary benefits. Although the economic benefits brought by nondifferential pricing of e-commerce company are less in the short term, the reputation benefits and social benefits can create greater economic benefits for the development of the company in the long term, which are the wealth of e-commerce company. The reputation benefits and social benefits brought by nondifferential pricing can be benefit for the more fair and equitable overall development environment for e-commerce.

This study systematically analyzes the model on the supervision of “big data killing” in the e-commerce company. However, the mechanism setting of the four-party game in the study has been simplified to a certain extent, and the strategy space needs to be more detailed and in-depth, which should be improved in the future. Moreover, because the simulation data were conducted under simulated conditions according to actual conditions, there may be some deviations in the effectiveness of players’ behavior analysis in the “big data killing” game. In the future, methods such as data mining will be used to collect big data, and empirical analysis of evolutionary game will be carried out, to improve the research on the participants behavior of “big data killing” in e-commerce transactions.

## Appendix

### A. Jacobian Matrix of Replicated Dynamical Systems

$$J = \begin{bmatrix} F'_x(x) & F'_y(x) & F'_z(x) & F'_r(x) \\ F'_x(y) & F'_y(y) & F'_z(y) & F'_r(y) \\ F'_x(z) & F'_y(z) & F'_z(z) & F'_r(z) \\ F'_x(r) & F'_y(r) & F'_z(r) & F'_r(r) \end{bmatrix}. \quad (\text{A.1})$$



Among them,

$$\begin{aligned}
F'_x(x) &= (1-2x)\{P_n - P_d - y\Delta P + y(M + I_e)[r + (1-r)\alpha] + (1-y)rI_e + zF\}, \\
F'_y(x) &= x(1-x)\{-\Delta P + (M + I_e)[r + (1-r)\alpha] - rI_e\}, \\
F'_z(x) &= x(1-x)F, \\
F'_r(x) &= x(1-x)[y(M + I_e)(1-\alpha) - yI_e], \\
F'_x(y) &= y(1-y)\{\Delta P - [M + (1-r)\alpha M] + (1-r)C_c - z(1-\beta)F\}, \\
F'_y(y) &= (1-2y)\{U_l - U_d - (1-x)\Delta P + (1-x)[M + (1-r)\alpha M] - (1-x)(1-r)C_c + (1-x)z(1-\beta)F\}, \\
F'_z(y) &= y(1-y)\{(1-x)(1-\beta)F\}, \\
F'_r(y) &= y(1-y)\{(1-x)(-\alpha)M\} + (1-x)C_c\}, \\
F'_x(z) &= z(1-z)\{-\beta F + rI_p + y(1-r)I_p\}, \\
F'_y(z) &= z(1-z)\{T_p - (1-x)(1-r)I_p\}, \\
F'_z(z) &= (1-2z)\{\beta F - C_p + yT_p - x\beta F - (1-x)rI_p - (1-x)y(1-r)I_p\}, \\
F'_r(z) &= z(1-z)\{-(1-x)I_p + (1-x)yI_p\}, \\
F'_x(r) &= r(1-r)\{R - I_e - N - (1-z)I_p + y\alpha[I_e + (1-z)I_p]\}, \\
F'_y(r) &= r(1-r)\{-(1-x)\alpha[I_e + (1-z)I_p]\}, \\
F'_z(r) &= r(1-r)\{-(1-x)I_p + (1-x)y\alpha I_p\}, \\
F'_r(r) &= (1-2r)\{-C_g + xR + (1-x)I_e + (1-x)(1-z)I_p + (1-x)N - (1-x)y\alpha[I_e + (1-z)I_p]\}.
\end{aligned} \tag{A.2}$$

## B. Jacobian Matrix of Equilibrium Points of Reproduced Dynamic Systems under Government Strict Supervision

The Jacobian matrix of the equilibrium point E1 (0, 0, 0, 1) is as follows:

$$J_1 = \begin{bmatrix} P_n - P_d + I_e & 0 & 0 & 0 \\ 0 & U_l - U_d - \Delta P + M & 0 & 0 \\ 0 & 0 & \beta F - C_p - I_p & 0 \\ 0 & 0 & 0 & -I_e + C_g - N - I_p \end{bmatrix}. \tag{B.1}$$

The Jacobian matrix of the equilibrium point E2 (1, 0, 0, 1) is as follows:

$$J_2 = \begin{bmatrix} -P_n + P_d - I_e & 0 & 0 & 0 \\ 0 & U_l - U_d & 0 & 0 \\ 0 & 0 & -C_p & 0 \\ 0 & 0 & 0 & C_g - R \end{bmatrix}. \tag{B.2}$$

The Jacobian matrix of the equilibrium point E3 (0, 1, 0, 1) is as follows:

$$J_3 = \begin{bmatrix} P_n - P_d + \Delta P + M + I_e & 0 & 0 & 0 \\ 0 & -U_l + U_d + \Delta P - M & 0 & 0 \\ 0 & 0 & \beta F - C_p + T_p - I_p & 0 \\ 0 & 0 & 0 & C_g - (1 - \alpha)(I_e + I_p) - N \end{bmatrix}. \quad (\text{B.3})$$

The Jacobian matrix of the equilibrium point E4 (0, 0, 1, 1) is as follows:

$$J_4 = \begin{bmatrix} P_n - P_d + I_e + F & 0 & 0 & 0 \\ 0 & U_l - U_d - \Delta P + M + (1 - \beta)F & 0 & 0 \\ 0 & 0 & -\beta F + C_p + I_p & 0 \\ 0 & 0 & 0 & C_g - I_e - N \end{bmatrix}. \quad (\text{B.4})$$

The Jacobian matrix of the equilibrium point E5 (1, 1, 0, 1) is as follows:

$$J_5 = \begin{bmatrix} -P_n + P_d + \Delta P - M - I_e & 0 & 0 & 0 \\ 0 & -U_l + U_d & 0 & 0 \\ 0 & 0 & -C_p + T_p & 0 \\ 0 & 0 & 0 & C_g - R \end{bmatrix}. \quad (\text{B.5})$$

The Jacobian matrix of the equilibrium point E6 (1, 0, 1, 1) is as follows:

$$J_6 = \begin{bmatrix} -P_n + P_d - I_e - F & 0 & 0 & 0 \\ 0 & U_l - U_d & 0 & 0 \\ 0 & 0 & C_p & 0 \\ 0 & 0 & 0 & C_g - R \end{bmatrix}. \quad (\text{B.6})$$

The Jacobian matrix of the equilibrium point E7 (0, 1, 1, 1) is as follows:

$$J_7 = \begin{bmatrix} P_n - P_d - \Delta P + M + I_e + F & 0 & 0 & 0 \\ 0 & -U_l + U_d + \Delta P - M - (1 - \beta)F & 0 & 0 \\ 0 & 0 & -\beta F + C_p - T_p + I_p & 0 \\ 0 & 0 & 0 & C_g - (1 - \alpha)I_e - N \end{bmatrix}. \quad (\text{B.7})$$

The Jacobian matrix of the equilibrium point E8 (1, 1, 1, 1) is as follows:

$$J_8 = \begin{bmatrix} -P_n + P_d + \Delta P - M - I_e - F & 0 & 0 & 0 \\ 0 & -U_l + U_d & 0 & 0 \\ 0 & 0 & C_p - T_p & 0 \\ 0 & 0 & 0 & C_g - R \end{bmatrix}. \quad (\text{B.8})$$

### C. Jacobian Matrix of Equilibrium Points of Reproduced Dynamic Systems under Government Loose Supervision

The Jacobian matrix of the equilibrium point E9 (0, 0, 0, 0) is as follows:

$$J_9 = \begin{bmatrix} P_n - P_d & 0 & 0 & 0 \\ 0 & U_l - U_d - \Delta P + (1 + \alpha)M - C_c & 0 & 0 \\ 0 & 0 & \beta F - C_p & 0 \\ 0 & 0 & 0 & C_g - I_e - I_p - N \end{bmatrix}. \quad (\text{C.1})$$

The Jacobian matrix of the equilibrium point E10 (1, 0, 0, 0) is as follows:

$$J_{10} = \begin{bmatrix} -P_n + P_d & 0 & 0 & 0 \\ 0 & U_l - U_d & 0 & 0 \\ 0 & 0 & \beta F - C_p & 0 \\ 0 & 0 & 0 & -C_g + I_e + I_p + N \end{bmatrix}. \quad (\text{C.2})$$

The Jacobian matrix of the equilibrium point E11 (0, 1, 0, 0) is as follows:

$$J_{11} = \begin{bmatrix} P_n - P_d - \Delta P + \alpha(M + I_e) & 0 & 0 & 0 \\ 0 & -U_l + U_d + \Delta P - (1 + \alpha)M + C_c & 0 & 0 \\ 0 & 0 & \beta F - C_p + T_p - I_p & 0 \\ 0 & 0 & 0 & -C_g + (1 - \alpha)(I_e + I_p) + N \end{bmatrix}. \quad (\text{C.3})$$

The Jacobian matrix of the equilibrium point E12 (0, 0, 1, 0) is as follows:

$$J_{12} = \begin{bmatrix} P_n - P_d + F & 0 & 0 & 0 \\ 0 & U_l - U_d - \Delta P + (1 + \alpha)M - C_c + (1 - \beta)F & 0 & 0 \\ 0 & 0 & -\beta F + C_p & 0 \\ 0 & 0 & 0 & -C_g + I_e + N \end{bmatrix}. \quad (C.4)$$

The Jacobian matrix of the equilibrium point E13 (1, 1, 0, 0) is as follows:

$$J_{13} = \begin{bmatrix} -P_n + P_d + \Delta P - \alpha(M + I_e) & 0 & 0 & 0 \\ 0 & -U_l + U_d & 0 & 0 \\ 0 & 0 & -C_p + T_p & 0 \\ 0 & 0 & 0 & -C_g + R \end{bmatrix}. \quad (C.5)$$

The Jacobian matrix of the equilibrium point E14 (1, 0, 1, 0) is as follows:

$$J_{14} = \begin{bmatrix} -P_n + P_d - F & 0 & 0 & 0 \\ 0 & U_l - U_d & 0 & 0 \\ 0 & 0 & C_p & 0 \\ 0 & 0 & 0 & -C_g + R \end{bmatrix}. \quad (C.6)$$

The Jacobian matrix of the equilibrium point E15 (0, 1, 1, 0) is as follows:

$$J_{15} = \begin{bmatrix} P_n - P_d + \Delta P + \alpha(M + I_e) + F & 0 & 0 & 0 \\ 0 & -U_l + U_d + \Delta P - (1 + \alpha)M + C_c - (1 - \beta)F & 0 & 0 \\ 0 & 0 & -\beta F + C_p - T_p + I_p & 0 \\ 0 & 0 & 0 & -C_g + (1 - \alpha)I_e + N \end{bmatrix}. \quad (C.7)$$

The Jacobian matrix of the equilibrium point E16 (1, 1, 1, 0) is as follows:

$$J_{16} = \begin{bmatrix} -P_n + P_d + \Delta P - \alpha(M + I_e) - F & 0 & 0 & 0 \\ 0 & -U_l + U_d & 0 & 0 \\ 0 & 0 & C_p - T_p & 0 \\ 0 & 0 & 0 & -C_g + R \end{bmatrix}. \quad (C.8)$$

## Data Availability

The data used to support the findings of the study are available from the corresponding author upon request.

## Disclosure

The funder has no role in study design, data collection and analysis, publication decisions, or manuscript preparation.

## Conflicts of Interest

The author declares no conflicts of interest.

## Acknowledgments

This work was supported by the Key Project Fund of Scientific Research of Shenyang University of Technology (no. ZDZRGD2020041) and Social Science Planning Fund of Liaoning Province (no. L19BGL029).

## References

- [1] M. Yu, "Can 'big data killing' be avoided? E-commerce adverse selection risk aversion," *Price Theory & Practice*, vol. 4, pp. 141–144, 2019.
- [2] A. Cavallo, *More Amazon Effects: Online Competition and Pricing Behaviors*, National Bureau of Economic Research, Cambridge, MA, USA, 2018.
- [3] G. Xing, F. Lu, and Z. Zhou, "Can the right of data portability govern big data: based price discrimination?" *Chinese Journal of Management Science*, pp. 1–16, 2022.
- [4] M. Wallmark, E. Greenberg, and D. Engels, "Consumer welfare and price discrimination: a fine line," *SMU Data Science Review*, vol. 1, no. 1, pp. 1–17, 2018.
- [5] M. J. Tremblay, "Pareto price discrimination," *Economics Letters*, vol. 183, Article ID 108559, 2019.
- [6] J. Moriarty, "Why online personalized pricing is unfair," *Ethics and Information Technology*, vol. 23, no. 3, pp. 495–503, 2021.
- [7] C. Cauffman, "Editorial: discrimination in online advertising," *Maastricht Journal of European and Comparative Law*, vol. 28, no. 3, pp. 283–286, 2021.
- [8] Y. Lu, Y. Qi, S. Qi, Y. Li, H. Song, and Y. Liu, "Say no to price discrimination: decentralized and automated incentives for price auditing in ride-hailing services," *IEEE Transactions on Mobile Computing*, vol. 21, no. 2, pp. 663–680, 2022.
- [9] D. Yu and X. Li, "Dynamic evolutionary game research of big data-based price discrimination between consumers and merchants," *Price Theory & Practice*, vol. 11, pp. 129–132, 2019.
- [10] F. Peiseler, A. Rasch, and S. Shekhar, "Imperfect information, algorithmic price discrimination, and collusion," *The Scandinavian Journal of Economics*, Article ID 12465, 2021.
- [11] S. Colombo, "Behavior- and characteristic-based price discrimination," *Journal of Economics and Management Strategy*, vol. 27, no. 2, pp. 237–250, 2018.
- [12] S. Ichihashi, "Online privacy and information disclosure by consumers," *The American Economic Review*, vol. 110, no. 2, pp. 569–595, 2020.
- [13] L. Lei, S. Gao, and R. Chen, "How to solve the problem of big data killing: evolutionary game in e-commerce market based on collaborative supervision of government and consumers," *Journal of Systems Management*, vol. 30, no. 4, pp. 664–675, 2021.
- [14] M. Botta and K. Wiedemann, "To discriminate or not to discriminate? Personalised pricing in online markets as exploitative abuse of dominance," *European Journal of Law and Economics*, vol. 50, no. 3, pp. 381–404, 2020.
- [15] T. Hupperich, D. Tatang, N. Wilkop, and T. Holz, "An empirical study on online price differentiation," in *Proceedings of the Eighth ACM Conference on Data and Application Security and Privacy*, pp. 76–83, New York, NY, USA, March 2018.
- [16] A. Gautier, A. Ittoo, and P. Van Cleynenbreugel, "AI algorithms, price discrimination and collusion: a technological, economic and legal perspective," *European Journal of Law and Economics*, vol. 50, no. 3, pp. 405–435, 2020.
- [17] T. Speicher, M. Ali, G. Venkatadri et al., "Potential for discrimination in online targeted advertising," *Journal of Machine Learning Research*, vol. 81, pp. 5–19, 2018.
- [18] J. Poort and F. Z. Borgesius, "Does everyone have a price? Understanding people's attitude towards online and offline price discrimination," *Internet Policy Review*, vol. 8, no. 1, pp. 1–20, 2019.
- [19] A. Bonatti and G. Cisternas, "Consumer scores and price discrimination," *The Review of Economic Studies*, vol. 87, no. 2, pp. 750–791, 2020.
- [20] M. Nuccio and M. Guerzoni, "Big data: hell or heaven? Digital platforms and market power in the data-driven economy," *Competition and Change*, vol. 23, no. 3, pp. 312–328, 2019.
- [21] A. Merikhnejadasl, K. Hridarzadeh, and Y. Mansourian, "Consumer responses to price discrimination: an experimental study," *Iranian Journal of Management Sciences*, vol. 14, no. 54, pp. 70–90, 2019.
- [22] E. Mauring, *Search and price Discrimination Online*, CEPR Discussion Papers, Article ID 15729, 2021.
- [23] D. Ngwe, K. J. Ferreira, and T. Teixeira, "The impact of increasing search frictions on online shopping behavior: evidence from a field experiment," *Journal of Marketing Research*, vol. 56, no. 6, pp. 944–959, 2019.
- [24] Y. Tang, R. R. Chen, and X. Guan, "Daily-deal market with consumer retention: price discrimination or quality differentiation," *Omega*, vol. 102, Article ID 102330, 2021.
- [25] A. Chandra and M. Lederman, "Revisiting the relationship between competition and price discrimination," *American Economic Journal: Microeconomics*, vol. 10, no. 2, pp. 190–224, 2018.
- [26] S. Hidir and N. Vellodi, "Privacy, personalization, and price discrimination," *Journal of the European Economic Association*, vol. 19, no. 2, pp. 1342–1363, 2021.
- [27] J. Huang, A. Mani, and Z. Wang, "The value of price discrimination in large random networks," in *Proceedings of the 2019 ACM Conference on Economics and Computation*, pp. 243–244, Hong Kong, Shenzhen, China, June 2019.
- [28] R. Zhang, J. Minjoon, and P. Sergio, "M-shopping Service Quality Dimensions and Their Effects on Customer Trust and Loyalty: An Empirical Study," *International Journal of Quality & Reliability Management*, 2021.
- [29] O. Bar-Gill, "Algorithmic price discrimination: When demand is a function of both preferences and (mis) perceptions," *University of Chicago Law Review*, vol. 86, pp. 18–32, 2018.
- [30] B. Wu, J. Cheng, and Y. Song, "Evolutionary game analysis of the phenomenon of deceiving acquaintances in e-commerce platform from the perspective of mental account," *Operations Research and Management Science*, vol. 29, no. 11, pp. 37–44, 2020.

- [31] C. Lindgren, S.-O. Daunfeldt, N. Rudholm, and S. Yella, "Is intertemporal price discrimination the cause of price dispersion in markets with low search costs?" *Applied Economics Letters*, vol. 28, no. 11, pp. 968–971, 2021.
- [32] J. A. Chevalier and A. K. Kashyap, "Best prices: price discrimination and consumer substitution," *American Economic Journal: Economic Policy*, vol. 11, no. 1, pp. 126–159, 2019.
- [33] C. Choe, S. King, and N. Matsushima, "Pricing with cookies: behavior-based price discrimination and spatial competition," *Management Science*, vol. 64, no. 12, pp. 5669–5687, 2018.
- [34] Z. Zhou, B. Wang, B. Gu et al., "Time-dependent pricing for bandwidth slicing under information asymmetry and price discrimination," *IEEE Transactions on Communications*, vol. 68, no. 11, pp. 6975–6989, 2020.
- [35] N. Fabra and M. Reguant, "A model of search with price discrimination," *European Economic Review*, vol. 129, Article ID 103571, 2020.
- [36] T. M. Rofin and B. Mahanty, "Impact of wholesale price discrimination on the profit of chain members under different channel power structures," *Journal of Revenue and Pricing Management*, vol. 20, no. 2, pp. 91–107, 2021.
- [37] A. Leibbrandt, "Behavioral constraints on price discrimination: experimental evidence on pricing and customer antagonism," *European Economic Review*, vol. 121, Article ID 103303, 2020.
- [38] K. Ritzberger and J. W. Weibull, "Evolutionary selection in normal-form games," *Econometrica*, vol. 63, no. 6, pp. 1371–1399, 1995.
- [39] R. Selten, "A note on evolutionarily stable strategies in asymmetric animal conflicts," *Journal of Theoretical Biology*, vol. 84, no. 1, pp. 93–101, 1980.
- [40] S. Sun and Q. Su, "Four-Party evolutionary game analysis of quality supervision of medical research reports in medical research reports in major epidemic period," *Chinese Journal of Management*, vol. 17, no. 9, pp. 1391–1401, 2020.



## Retraction

# Retracted: A Mode of Intelligent Equipment Monitoring Optimization Based on Dynamic Programming Algorithm

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] Z. Zheng and C. Jiang, "A Mode of Intelligent Equipment Monitoring Optimization Based on Dynamic Programming Algorithm," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 1569428, 11 pages, 2022.

## Research Article

# A Mode of Intelligent Equipment Monitoring Optimization Based on Dynamic Programming Algorithm

Zhilei Zheng  and Chuan Jiang 

*Department of Physical Education, Shandong University, Weihai 264200, China*

Correspondence should be addressed to Zhilei Zheng; 200869800016@sdu.edu.cn

Received 17 January 2022; Revised 13 February 2022; Accepted 28 February 2022; Published 18 March 2022

Academic Editor: Daqing Gong

Copyright © 2022 Zhilei Zheng and Chuan Jiang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the continuous development of the national economy and scientific productivity, urban construction and people's living standards are also getting higher and higher. Although people enjoy increasingly convenient life, the demand for intelligence is getting higher and higher. Digital intelligent equipment has the functions of data collection, calculation and analysis, diagnostic and early warning, and communication functions. Analyze the status quo and existing problems of the development of intelligent equipment, as well as analyze and research key monitoring technologies in the use and development of digital intelligent equipment and provide optimal solutions for intelligent equipment hardware development requirements, software development, and model algorithms. Intelligent equipment monitoring is related to all aspects of people's livelihood, and its intelligent development is related to the public role in this field in the future. Accurate results of monitoring can provide data support for schools, research institutions, the public, and the government. At the same time, it is also an important basis for formulating social policies. At present, the commonly used monitoring method usually adopts time series algorithm. Through literature review, it is found that the algorithm has the problem of distortion of correct data, which affects the accuracy of monitoring results. Based on the above reasons, this article combines the wavelet function with the planning algorithm and proposes a dynamic programming algorithm, which removes the redundant monitoring data in turn and clusters the distortion monitoring data with the wavelet function, which improves the accuracy and computational efficiency of the algorithm and gives full play to the monitoring of intelligence. The simulation results using MATLAB show that the planning algorithm can eliminate 90% of redundant monitoring data and improve the extraction rate of characteristic monitoring data. At the same time, the accuracy of the planning algorithm reaches 95%, and the calculation time is less than 25 s, which is better than the static planning algorithm. Therefore, the dynamic programming algorithm can better utilize the intelligence, convenience, and efficiency of the equipment to optimize the monitoring model.

## 1. Introduction

Intelligent devices have been widely used in engineering technology, industrial production, social services, medicine, and human health, and they have had an important impact. Intelligent equipment monitoring is an important part of social public health. It is not only the guarantee of Chinese public health but also the basis for the improvement of national quality [1]. The integration of intelligent equipment and modern technology not only enhances the intelligence of monitoring but also realizes its own function optimization. Continuous and dynamic intelligent equipment

monitoring and analysis of the public can help the government make health decisions and improve the disease resistance and physical quality of the public Beuchat et al. [2]. In the past, the public monitoring method was mainly planning algorithm. Although this algorithm can monitor the general health, it cannot carry out dynamic analysis, and the analysis results have some limitations. In addition, there are a lot of redundant data in intelligent equipment monitoring, which affects the monitoring results of college students [3]. At present, there are many researches on intelligent equipment device monitoring of college students, but the redundancy of monitoring data is also high, as shown

in Figure 1. Therefore, the application of dynamic programming method in the optimal control of intelligent equipment has very important theoretical significance. At the same time, the large-scale application of intelligent devices needs more effective algorithms as support, so this study can provide more practical cases, which has very important practical significance.

Some scholars such as Fortune et al. [4] also believe that the increasing intelligent characteristic data in the static programming algorithm will improve the intelligence of monitoring but reduce the accuracy of monitoring. Therefore, a dynamic programming algorithm is constructed and revised to improve the accuracy of the results [5]. As a processing method of comprehensive function, wavelet function can subcontract the content data into fragmented data groups, improve the processing capacity of initial data, and is widely used in the optimization of a large number of intelligent devices. However, it is not used in dynamic programming algorithm at present. Therefore, this article attempts to apply it to the dynamic programming algorithm and construct a new intelligent device optimization model, which aims to improve the data processing efficiency of intelligent devices and promote the development of intelligent devices. Dynamic programming algorithm is developed on the basis of static programming algorithm. It is to add time variables to the original algorithm and continuously combine the static programming algorithm. However, the fundamental difference between dynamic algorithm and static algorithm is to increase the coupling between different indexes and make the calculation results more continuous. At the same time, the improvement of coupling also increases the complexity of calculation and the incidence of redundant data. Based on the research of scholars at home and abroad, the integration of wavelet function and dynamic programming algorithm can use the subcontracting method of wavelet function to increase the data processing capacity of dynamic programming algorithm, reduce the occurrence rate of local extremum, and improve the accuracy of calculation results. Dynamic programming algorithm can continuously track intelligent devices without being affected by the increase of data volume. Therefore, the integration of wavelet function and dynamic programming algorithm has the advantage of comprehensive analysis and improves the processing capacity of calculation data. Based on the above reasons, this article combines the discrete wavelet function with the planning algorithm to construct a dynamic planning algorithm to optimize the health monitoring results of College Students' wearing devices [6]. Taking college students' physical health monitoring as an analysis case, this article improves the accuracy of the results, so as to better carry out intelligent health monitoring.

## 2. Algorithm Description of Intelligent Health Monitoring Devices

The key of intelligent equipment monitoring is to quantify the relevant indicators in intelligent equipment devices, so as to lay a mathematical foundation for dynamic monitoring. Among them, height, weight, 50-meter run, 800-meter run

(girls), and 1000-meter run (boys) have more wearing equipment [7], whereas there are relatively few data on vital capacity, standing long jump, sitting body flexion, sit ups, step test and pull-up, but the algorithm description of health monitoring data are required.

### 2.1. Intelligent Health Monitoring Process of College Students.

Intelligent health monitoring of intelligent equipment devices includes three aspects: the data structure  $x_i$  in sports intelligent equipment devices, the impact of intelligence on health monitoring  $x_j$ , and the dynamics  $x_k$  of health monitoring [8]. The data structure in the intelligent equipment device also includes the intelligent equipment device  $x_{i1}$ , the proportion of different sports  $x_{i2}$ , and the degree of cooperation between different sports  $x_{i3}$ ; similarly, the impact of intelligence on health monitoring include the promotion degree of intelligent technology to health monitoring  $x_{j1}$ , the integration degree of intelligent technology and health monitoring  $x_{j2}$ , and the promotion level of intelligent technology to health monitoring  $x_{j3}$ ; The dynamics of health monitoring include cooperation degree between cardio-pulmonary monitoring, blood oxygen monitoring, blood pressure monitoring, and exercise monitoring  $x_{k1}$ , cooperation between devices in different forms of intelligent technology  $x_{k2}$ , and communication cooperation between different intelligent technologies  $x_{k3}$ . According to the above analysis, college students' intelligent equipment monitoring involves many aspects, mainly including: height, weight, vital capacity, standing long jump, sitting forward flexion, 50 m running, sit ups, step test, pull up, 800 m running (girls), and 1000 m running (boys). At the same time, the collected monitoring data are massive (cloud monitoring data, a large number of applications of intelligent devices) and complex (there are a large number of unstructured monitoring data), which greatly reduces the accuracy of intelligent equipment monitoring in sports and cannot accurately reflect the intelligence of intelligent equipment devices. Because the mass and complexity are the inevitable trend of the development of intelligent equipment monitoring, a large amount of sports data of college students should be collected [9].

### 2.2. Data Flow Description of College Students' Intelligent Equipment Monitoring.

The stable wavelet function extracts the redundant data in college students' sports, makes discrete analysis on the intelligent equipment devices, and obtains the comprehensive monitoring data, so as to arrange the data in order to improve the accuracy of monitoring [10].

At the same time, the stable wavelet function uses discrete extraction to ensure the integrity of monitoring data in college students' sports and complete the single-phase characteristics of monitoring data in order. The specific monitoring data flow is described in Figure 2:

- (1) Assuming that the intelligent equipment monitoring result is  $A$  and any intelligent equipment device  $A_i = \{a_1, a_2, \dots, a_n\}$ , the relationship between  $A$  and each input monitoring data as follows:

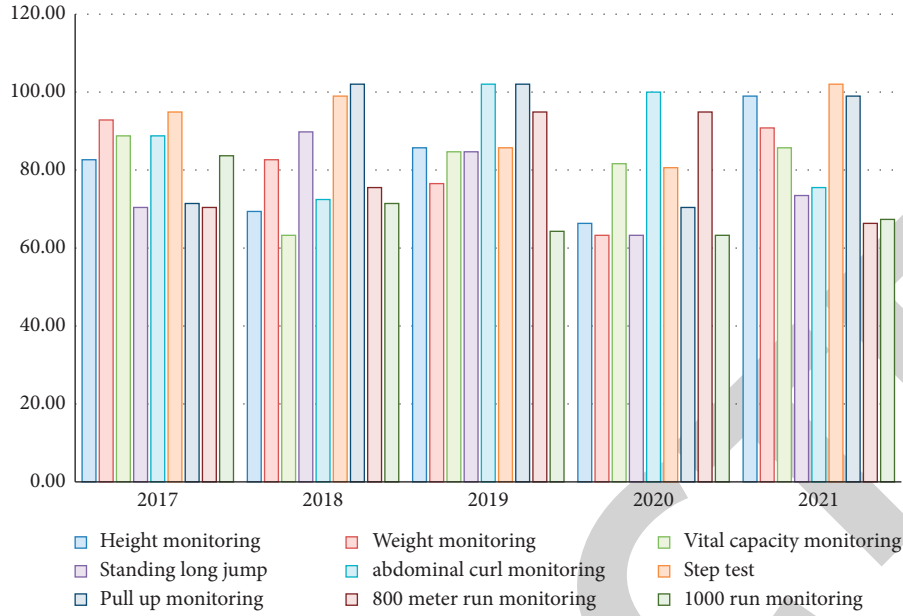


FIGURE 1: The monitoring proportion of intelligent equipment devices in different sports events from 2017 to 2021.

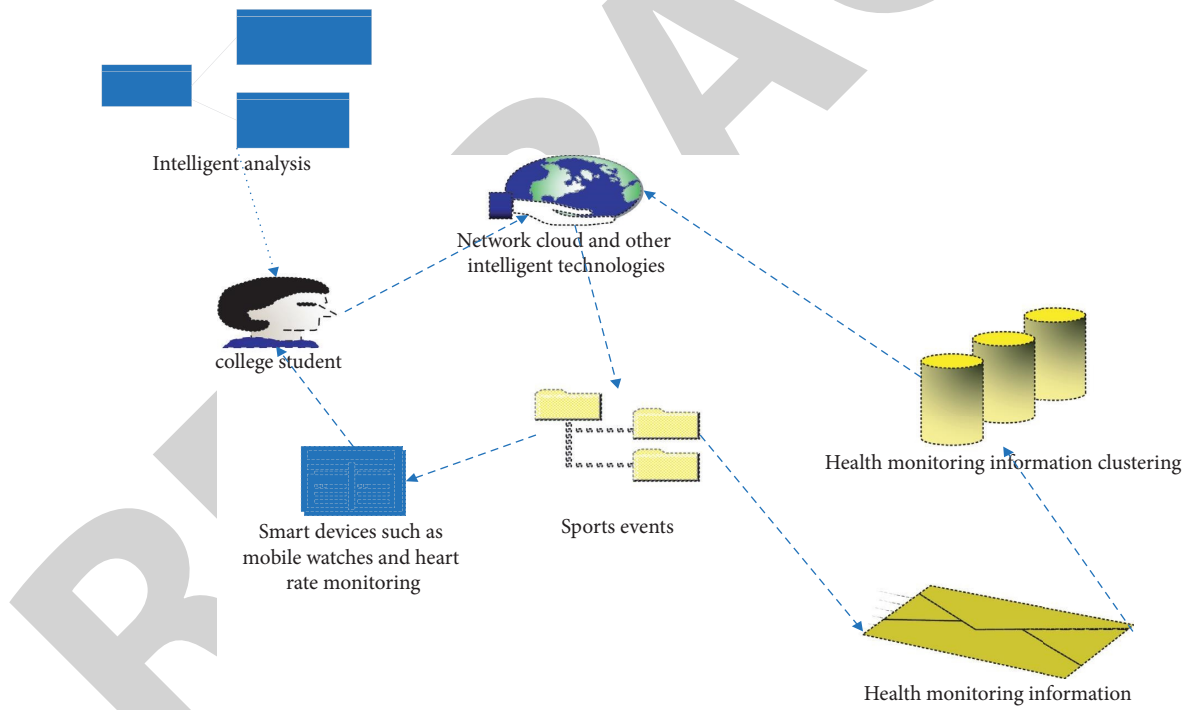


FIGURE 2: The intelligent equipment monitoring of intelligent equipment devices on College Students' sports events.

$$\sum_l^n A_l \overset{TS(\cdot)}{\underset{f(\cdot)}{\leftarrow}} \oint_n \overset{k\text{-means}}{\prod_{\text{indevie } i} \sum_i^n x_i \overset{g(\bullet)}{\rightleftharpoons} \sum_j^n x_j \overset{g(\bullet)}{\rightleftharpoons} \sum_k^n x_k}, \quad (1)$$

where  $i, j, k$  are natural numbers,  $TS(\cdot)$  is the dynamic programming function,  $f(\cdot)$  is the stable wavelet function,  $G(\cdot)$  is the forward function between different input indexes, and  $-G(\cdot)$  is the reverse function to complete the dynamic monitoring.

(2) Suppose that the monitoring result  $a_l$  of any sports event in intelligent health monitoring and the input  $z$  in the planning algorithm (data structure  $x_i$  in intelligent equipment devices [11], impact of intelligent technology on health monitoring  $x_j$ , and dynamic  $x_k$  of health monitoring),  $p$  is the proportion of monitoring data (structured monitoring data  $>70\%$ , semistructured monitoring data  $>70\%$ , unstructured monitoring data  $>70\%$ ),  $q$  is the processing method of monitoring data distortion

(reconstruction = 1, coefficient order = 2, quantification = 3, eigenvalue = 4, clustering = 4), then  $c_l$  is described as  $\text{Inc}_l^{o,p,q}$ ,  $o \in (1, 2, \dots, n)$ ,  $p \in (1, 2, 3, 4)$ ,  $q \in (1, 2, 3, 4)$ .

Among them, the logarithm in  $\ln(\cdot)$  of  $c_l$  is used to avoid  $|\infty|$  or extreme value 0, so as to ensure the effectiveness of the calculation results. The monitoring results of any sports are shown in Figure 2.

- (3) The data structure  $x_i$  in devices, the impact of intelligence on health monitoring  $x_j$ , and the dynamic  $x_k$  of health monitoring adopt the fusion function  $f(x)$  and propose the redundant monitoring data in college students' sports. The calculation formula of the function is as follows:

$$\varphi(x) = \frac{\left[ \left( \sum_{i=1}^n \alpha_1 x_i + \sum_{j=1}^n \alpha_2 x_j + \sum_{k=1}^n \alpha_3 x_k \right) + \xi \right]}{\sum_{i,j,k=1}^n (x_i + x_j + x_k)}, \quad (2)$$

where,  $\alpha_1$ ,  $\alpha_2$ , and  $\alpha_3$  are the weight coefficients of intelligent health monitoring equipment, college students, and sports, respectively, and  $\xi$  is the adjustment error. The weight coefficient  $\xi$  is calculated from the data published by the National College Students' intelligent health monitoring organization [12].

- (4) The health monitoring data of dynamic planning is mainly based on the statistical data of the National Center for Disease Control and Prevention. The intelligent health monitoring time is calculated as  $T$ , the evaluation results  $\theta$ , and the simplification rate of health monitoring data = (monitoring data before processing - monitoring data after processing) / (the total number of monitoring data obtained (100 m running, standing long jump, etc.) \* 100%). The specific formula is as follows:

$$\iota = \frac{\|H(\cdot) - G_{i,j,k}\|}{\|G_i + G_j + G_k\|} \cdot \begin{pmatrix} \alpha_i & 0 & 0 \\ 0 & \alpha_j & 0 \\ 0 & 0 & \alpha_k \end{pmatrix} \cdot 100\%, \quad (3)$$

where  $G_{i,j,k}$  is the intelligent monitoring data of  $x_i$ ,  $x_j$ , and  $x_k$ ,  $G_i$ ,  $G_j$ , and  $G_k$  is the total amount of  $x_i$ ,  $x_j$ , and  $x_k$ ,  $H(\cdot)$  is the processed intelligent monitoring data,

and  $\begin{pmatrix} \alpha_i & 0 & 0 \\ 0 & \alpha_j & 0 \\ 0 & 0 & \alpha_k \end{pmatrix}$  is the weight of the intelligent monitoring data in  $x_i$ ,  $x_j$ , and  $x_k$ .

- (5) In order to improve the complex data (structured, semi-structured and unstructured) in college students' sports, realize the "intelligence" of the transformation between health monitoring data and reduce the error of initial monitoring data, subcontracting can be used for processing. Because the attribute of intelligent monitoring data is relatively simple, the Euclidean distance can be used for

wavelet packet. First set the range of structured, unstructured, and semistructured monitoring data, and the formula is as follows:

$$|S| = \frac{c_l}{\left( \sum_{l=0}^n c_l + \mu^2 \right)}, \quad (4)$$

where  $|S|$  is the Euclidean distance of each monitoring data;  $c_l$  is the intelligent;  $c_0$  is the initial cluster value;  $\mu$  is the allowable error of clustering and is set by the power industry in the early stage of calculation,  $\sum_{l=0}^n c_l$  is the set of  $c_l$ . According to formula (1) to formula (4) steps to calculate the conclusion, the dynamic programming algorithm uses the judgment matrix of wavelet function to subcontract the overall data. At the same time, sort the subcontracted data to form an orderly data flow. On this basis, the data processing capacity of dynamic programming method decreases exponentially. At the same time, the continuous tracking of dynamic programming algorithm increases the total amount of data, but the amount of data processing can be further improved by setting the subcontracting value. Therefore, the advantages of wavelet function and dynamic programming algorithm in data processing are very obvious.

### 3. Build an Intelligent Health Monitoring Model Based on Dynamic Programming Algorithm

3.1. Construct the Relationship among College Students' Data Structure, Health Monitoring Impact, and Health Monitoring Dynamics. Intelligent health monitoring not only considers the monitoring accuracy of  $x_i$ ,  $x_j$ , and  $x_k$  but also considers the synergy among them. Therefore, the relationship operator between  $x_i$ ,  $x_j$ , and  $x_k$  should be constructed. It is assumed that the relationship between the three is divided into local fusion relationship ( $x_i$  and  $x_j$ ,  $x_i$  and  $x_k$ , and  $x_j$  and  $x_k$ ) and overall fusion relationship ( $x_i, x_j, x_k$ ), which are expressed by  $P_c$  and  $P_m$ , respectively. Then, the calculation formula is as follows:

$$P_c = P_m \begin{cases} P_{c_0} \\ \varphi(x) \left[ \frac{(P_{c_{x_i,j,k}} - P_{\min})}{\sum P_{c_{x_i,j,k}} - P_{\max}} \right], \\ P_m \approx 0 \text{ or } 1 \\ P_{\max} \langle 1, P_m \rangle 1 \end{cases} \quad (5)$$

where,  $P_{c_0}$  is local fusion, all fusion,  $P_{\min}$  and  $P_{\max}$  are the minimum  $P_c$  and  $P_m$  maximum values, which are set by local universities, and  $\varphi(x)$  is the fusion degree function of the above analysis, the value is between 0 and 1.

3.2. *Construct the Accuracy Model of College Students' Intelligent Health Monitoring.* Accuracy is the key index of intelligent health monitoring [13], and it is also the main

$$Inc_l^{o,p,q} = \begin{cases} 0 < Inc_l^{o,p,q} < 1 \\ Inc_l^{o,p,q} < \left\| \max [Inc_l^{n,n,n}] \right\| \\ \int_{I,j}^n \left( \sum_{I=1}^n \max (Inc_l^{o,p,q}) \right) \begin{matrix} \gg \\ \ll \end{matrix} \min (Inc_l^{o,p,q}) \end{cases} |tP_c \neq P_m, P_{\max} < 1, \text{ and } P_m > 1 \rangle \quad (6)$$

Since the accuracy of  $x_i, x_j$  and  $x_k$  is  $0 < Inc_l^{o,p,q} < 1$ ,  $0 < Inc_l^{o,p,q} < 1$  are required to shorten the range of  $\max (Inc_l^{o,p,q}) \gg \min (Inc_l^{o,p,q})$  calculation accuracy. At the same time, the fusion error of the calculation results should be reduced, so  $\prod \varphi(x) \lim_{l \rightarrow \infty} \oint_l \sin(\alpha_{1,2,3}) 0 \sim 0.1$  should be between 0 and 0.1.

3.3. *Calculation Steps of Operators in Intelligent Health Monitoring of College Students' Intelligent Equipment Devices*

- (1) Build an intelligent health monitoring data set,  $C_l = \{c_1, c_2, \dots, c_i\}$  and eliminate the intelligent monitoring data with digitization degree of  $< 60\%$ . Determine the weight coefficient and adjustment error according to the digital Economic Yearbook of power industry from 2015 to 2020;
- (2) The monitoring data  $c_l$  processed by wavelet function are processed structurally [14], semistructurally and unstructured to realize the "cleaning" of intelligent monitoring data.
- (3) Judge the fusion relationship of the monitoring data after "cleaning" and obtain the values of  $P_c$  and  $P_m$  of local and overall fusion relationships.
- (4) The iterative analysis of steps 1~4 is carried out using MATLAB software, and the accuracy, calculation time, and effectiveness of the calculation results are output. The result are shown in Figure 3.

## 4. The the Case Analysis of Intelligent Health Monitoring Based on Dynamic Programming Algorithm

4.1. *The Actual Case Analysis.* Taking college students in various regions as the research object, analyze the data structure  $x_i$  in intelligent equipment devices, the impact of intelligence on health monitoring  $x_j$ , and the dynamics of health monitoring  $x_k$  as the input indicators. Local universities and national CDC provide  $c_0 = 22.3$ ,  $\mu = 0.02$ ,  $\alpha_0 = 0.41$ , 100 iterations, and  $P_{\min} = 0.23$ ,  $P_{\max} = 0.99$ . The specific monitoring data are shown in Table 1.

purpose of dynamic programming algorithm optimization. Therefore, the accuracy operator should be constructed. The specific formula is as follows:

4.2. *The Comparison of Data Processing Capacity of College Students' Intelligent Health Monitoring.* The dynamic programming algorithm uses wavelet packet to preprocess the intelligent monitoring data [15], eliminate redundant monitoring data, and extract eigenvalues. The processing result is better than the static programming algorithm. The results are shown in Table 2.

It can be seen from Table 2 that the dynamic programming algorithm is consistent with the actual demand, while the monitoring data processing capacity of the static programming algorithm is quite different from the actual demand and does not change along the change of the actual demand curve. In the range of 0~25 iterations, there is no significant difference in the processing capacity of monitoring data between the two algorithms, but there are significant differences in the range of 25~50, 50~75 and 75~100 iterations, and the dynamic programming algorithm is better than the static programming algorithm, which is consistent with the results of relevant studies Wang [16]. With the increase of the number of iterations, the amount of monitoring data processing increases. The specific results are shown in Figure 4.

It can be seen from Figure 4 that in the monitoring stage of 0.6 years, the change range of dynamic programming algorithm is significantly less than that of static detection algorithm. Moreover, the processing capacity of dynamic programming algorithm is higher than the total data processing capacity. Therefore, the amount of data processing in dynamic programming algorithm is more reasonable. At same time, it can be seen from Figure 4 that the data processing capacity of dynamic programming is relatively stable without significant fluctuation, while the fluctuation of static programming algorithm is relatively large in the early stage and tends to be flat in the later stage, so the processing effect of dynamic programming is better, which is consistent with the results of relevant studies [17]. The comparison of the actual data processing capacity of the two methods is shown in Figure 5.

As can be seen from Figure 5, the area of dynamic programming algorithm is significantly larger than that of static programming algorithm. Although the height of static programming algorithm is higher than that of dynamic algorithm, the overall area of dynamic programming algorithm is larger.



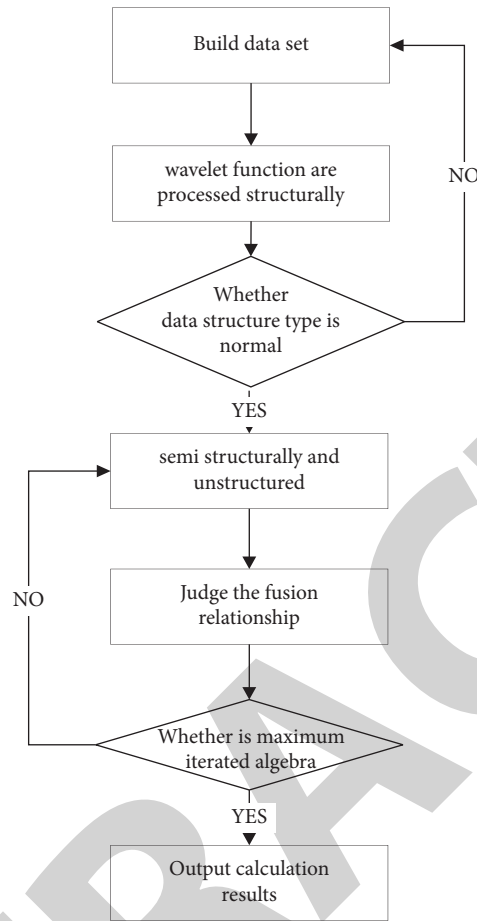


FIGURE 3: The Calculation flow of dynamic programming algorithm.

TABLE 1: The data sources of intelligent health monitoring of college students in actual cases.

Attribution area	Collection type	Collect content
Sports colleges and universities in Northeast China	ECG function monitoring, sleep quality, circulatory system	Big data, intelligent devices, cloud computing
Sports colleges and universities in North China	Subhealth, family history	Virtual simulation, smart client data collection
Physical education colleges and universities in Northwest China	Reproductive health, adolescent psychology, and blood health of female students	Intelligent device, mobile watch
Sports colleges and universities in Southeast China	Major genetic diseases and epidemic surveillance	Mobile hub, network platform
National Center for disease control and prevention	Prevention of sudden and serious diseases	Mobile phone, Big data
Private health monitoring organization	Sub ealth, blood pressure, blood lipid	University network platform
World health organization	Epidemic diseases, major diseases, and sudden diseases	QQ chatting software, VR game, Bluetooth headset big data,

Note: the monitoring data comes from the data of the National Center for Disease Control and prevention from 2015 to 2020. The choice of collection type and attribution area is based on the relevant literature at home and abroad and the actual survey results.

It can be seen from Figure 5 that the unit processing data of dynamic programming algorithm is significantly higher than that of static programming algorithm, which also fully explains the phenomenon that the processing process of dynamic programming algorithm is relatively stable. The main reason for the above problems is that the dynamic programming algorithm uses wavelet function for data subcontracting to reduce the complexity of data structure.

**4.3. The Calculation Accuracy.** The dynamic programming algorithm is used to analyze the data structure  $x_i$  in intelligent equipment devices, the impact of intelligence on health monitoring  $x_j$ , and the results of the dynamic  $x_k$  of health monitoring. The results are shown in Table 3.

It can be seen from Table 3 that the data structure  $x_i$  in intelligent equipment devices of dynamic programming algorithm, the impact of intelligence on health monitoring

TABLE 2: The average processing capacity of monitoring data of different algorithms.

Different methods	Iteration number range (times)	Average monitoring data processing capacity (km)
Dynamic programming algorithm	0~25	32.32 ± 2.32**
Standing long jump, 50-m sprint, 1000-vm long run, basketball	25~50	48.62 ± 3.28*
	50~75	65.28 ± 4.02*
	75~100	82.23 ± 1.23*
<b>Total processing capacity</b>		<b>228.45 ± 3.22*</b>
Static programming algorithm	0~25	22.37 ± 1.32
Standing long jump, 50-m sprint, 1000-m long run, basketball	25~50	40.31 ± 4.16
	50~75	55.08 ± 4.32
	75~100	72.13 ± 1.82
<b>Total processing capacity</b>		<b>189.89 ± 4.12</b>

Compared with static programming algorithm, \* $P < 0.05$ , \*\* $P > 0.05$ .

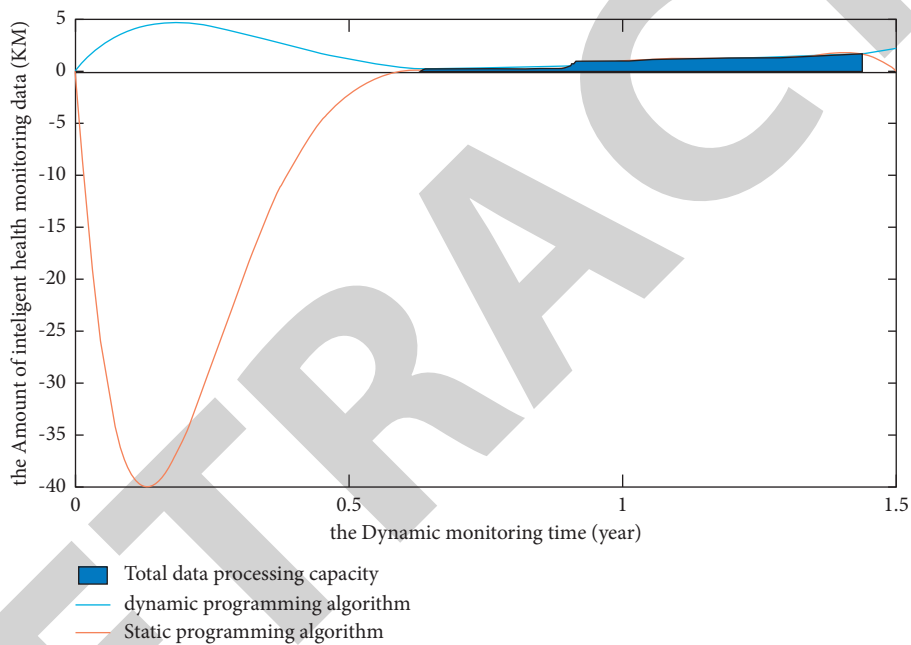


FIGURE 4: The processing capacity of intelligent health monitoring data at different times.

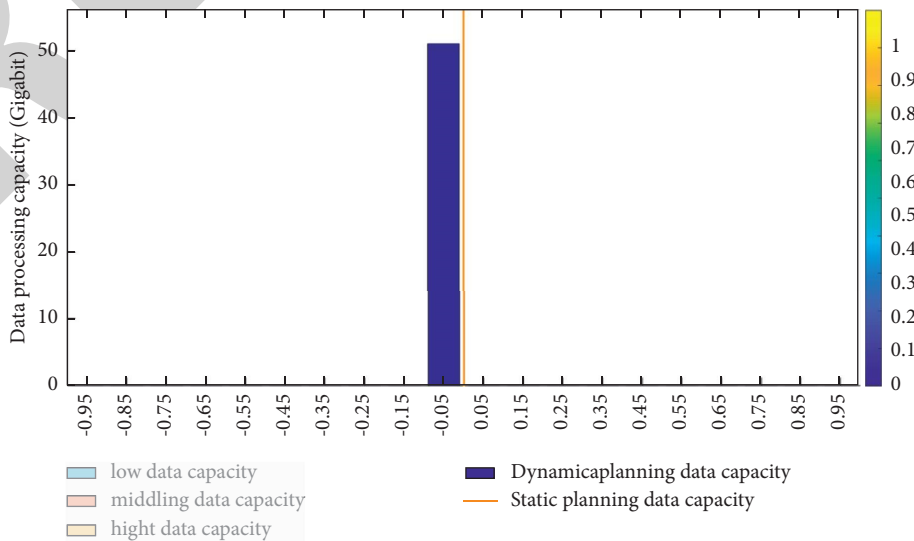


FIGURE 5: The data capacity of intelligent health monitoring at different times.

TABLE 3: The accuracy of different algorithms.

The different methods	The evaluating indicator	The accuracy rate (%)
The dynamic programming algorithm	$x_i$	98.23*
	$x_j$	99.23*
	$x_k$	97.62*
The total accuracy		98.32
The static programming algorithm	$x_i$	92.32
	$x_j$	96.23
	$x_k$	95.72
The total accuracy		95.98

Compared with static programming algorithm, \* $P < 0.05$ ; The value range is 0~100.

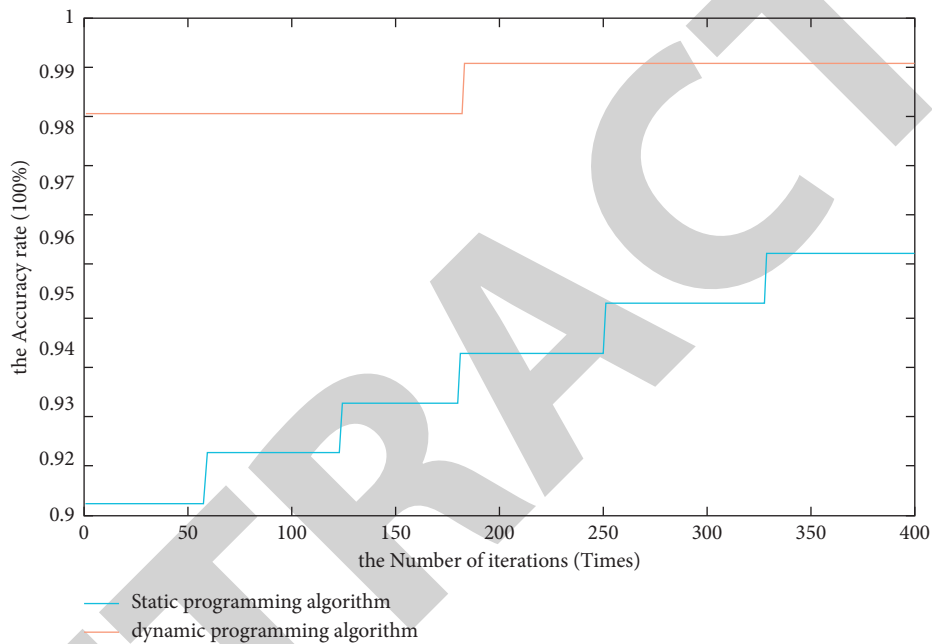


FIGURE 6: The calculation accuracy of different algorithms.

TABLE 4: The calculation time of different algorithms.

The different methods	The calculation time (seconds)
The dynamic programming algorithm	82.32
The static programming algorithm	81.22

Note: the value range is 0~100.

$x_j$ , and the dynamic  $x_k$  accuracy and overall accuracy of health monitoring are better than those of static programming algorithm. It mainly understands that the planning algorithm integrates the weight coefficient and adjustment error, which is consistent with the results of relevant studies [18]. The above three indicators show small error changes in the dynamic planning algorithm, so as to avoid the interference of redundant monitoring data in the calculation and improve the accuracy of calculation. The results are shown in Figure 6.

As can be seen from Figure 6, the accuracy of dynamic programming algorithm is always higher than that of static algorithm, and the overall fluctuation range is small. The accuracy of dynamic programming algorithm is between 98% and 99%, whereas that of static

programming algorithm is between 90% and 95%. Moreover, the stability of dynamic programming algorithm is better, which is significantly higher than that of static programming algorithm. The results are consistent with relevant domestic studies. The results are shown in Table 4.

**4.4. The Time of Intelligent Health Monitoring.** It can be seen from Table 4 that the health monitoring processing time of dynamic planning algorithm is significantly less than that of static planning algorithm, and the whole health monitoring processing time is less than 25 s. According to the line segment slopes of the two algorithms, the health monitoring data of static planning show ups and downs, whereas the

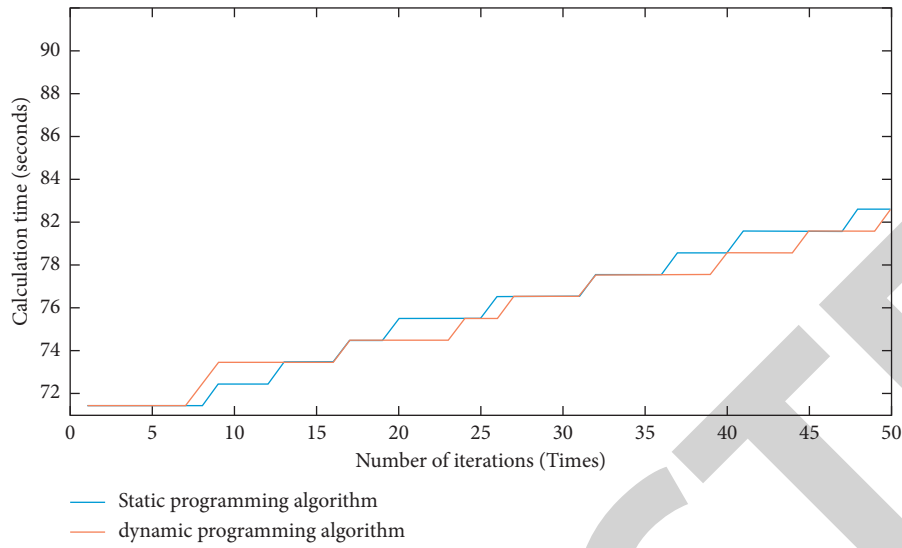


FIGURE 7: The calculation time of different algorithms.

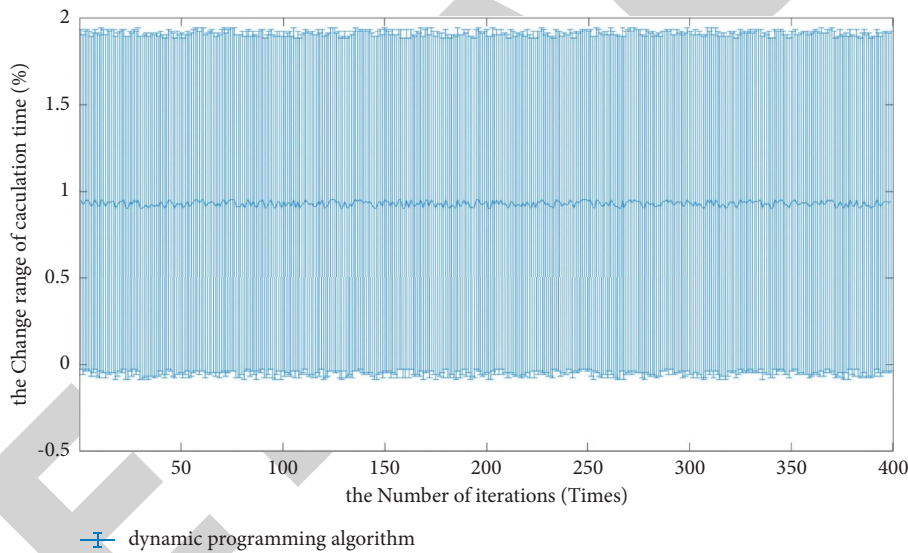


FIGURE 8: The change range of dynamic programming algorithm calculates time.

health monitoring data of dynamic planning change smoothly and the slope change range is small. This shows that the dynamic programming algorithm is better than the static programming algorithm, which can meet the requirements of intelligent health monitoring and also meet the standards of intelligent health monitoring by the ministry of health and local universities in 2021. The results are shown in Table 4.

As can be seen from Figure 7, the calculation time of dynamic programming method is always lower than that of static programming method, and the overall calculation time is almost the same. Therefore, the calculation time of dynamic programming is better. The variation range of

calculation time of different algorithms is shown in Figures 8 and 9.

As can be seen from Figures 7 and 8, the variation range of calculation time of dynamic programming algorithm is significantly lower than that of static programming algorithm. This fully explains the problem of stable calculation time of dynamic programming algorithm and also shows that the data processing effect of dynamic programming algorithm is better. The variation range of dynamic programming algorithm is 0.9%, while that of static programming algorithm is 100%. Therefore, the result of dynamic programming algorithm is more stable and can be used in data processing of high-end intelligent devices.

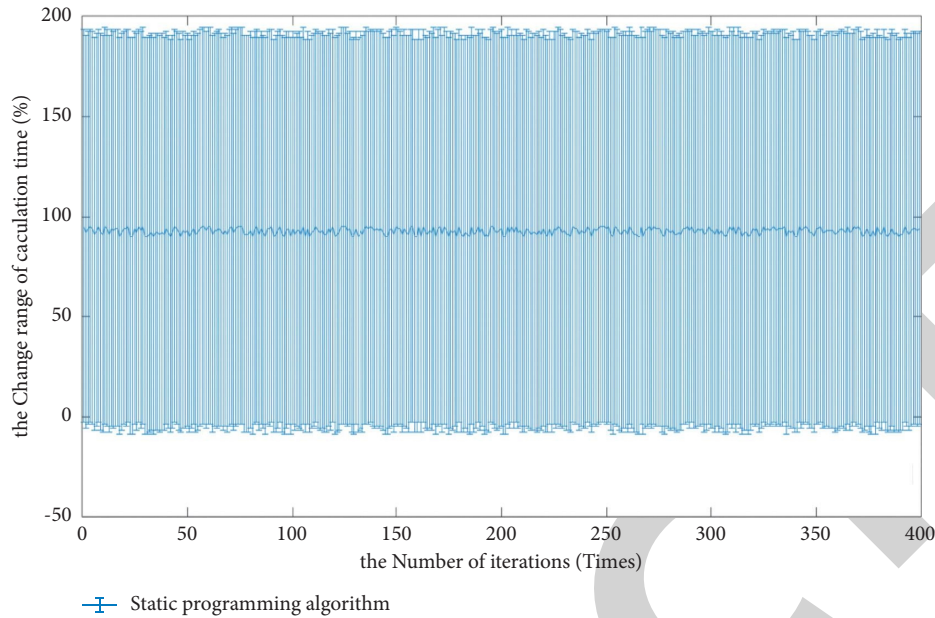


FIGURE 9: The change range of static programming algorithm calculates time.

## 5. Conclusion

By measuring the physical fitness of college students, evaluating the physical status and the effect of physical exercise, improving and urging the effective mechanism of college students to participate in physical activities, scientifically guiding college students to carry out physical activities, so as to continuously enhance the physical fitness of college students, thus also providing effective national physical fitness monitoring data can be done. Intellectualization is the requirement of the ministry of health and local colleges and universities for college students in 2020, and it is also the direction of future development and reform of colleges and universities [19]. Since 2010, the intelligent development of College Students' monitoring has gradually increased. However, due to the large amount of monitoring data involved in equipment public, relevant specifications, and other links, the previous static planning Bayesian and genetic algorithms have achieved accurate evaluation, which affects the implementation of later improvement measures [20]. In this article, wavelet function and static programming method are combined to build a dynamic programming algorithm model to evaluate the intelligent health monitoring. The results show that the following: (1) intelligent wearable devices are popular among contemporary college students and play an important supporting role in college students' intelligent health monitoring. At the same time, due to the great pressure of study and work, college students are not only the main subhealth population in society but also the key health monitoring object. The dynamic programming algorithm constructed in this article can continuously monitor the health of college students and has the advantages of continuous and complete monitoring data. Dynamic programming monitoring can provide redundant data and solve the problem of data distortion in static programming. (2) The constructed model can effectively eliminate redundant monitoring data and

accurately extract eigenvalues, which is consistent with the actual intelligent health monitoring requirements, and the calculation time is less than 25 s, and the evaluation process is stable. (3) Compared with static planning, the model is superior to static planning in terms of calculation accuracy. The accuracy of dynamic programming algorithm is between 98% and 99%, while that of static programming algorithm is between 90% and 95%. Moreover, the stability of dynamic programming algorithm is better, which is significantly higher than that of static programming algorithm. The results are consistent with relevant domestic studies. The monitoring data processing capacity and calculation time meets the standards of intelligent health monitoring by the ministry of health and local universities in 2021. The variation range of dynamic programming algorithm is 0.9%, while that of static programming algorithm is 100%. Therefore, the result of dynamic programming algorithm is more stable and can be used in data processing of high-end intelligent devices. Of course, this article has less research on the transmission technology and security technology of intelligent devices, and these contents are the guarantee of intelligence, so we will focus on the above two aspects in the future.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported by Shandong University, Weihai.

## Research Article

# A Random Forest Method for Identifying the Effectiveness of Innovation Factor Allocation

Mo Xu <sup>1,2</sup>, Yawei Qi <sup>3</sup>, Changqi Tao <sup>4</sup>, and Shangfeng Zhang <sup>1,2</sup>

<sup>1</sup>School of Statistics and Mathematics, Zhejiang Gongshang University, Hangzhou 310018, China

<sup>2</sup>Collaborative Innovation Center of Statistical Data Engineering Technology & Application, Zhejiang Gongshang University, Hangzhou 310018, China

<sup>3</sup>School of Information Management, Jiangxi University of Finance and Economics, Nanchang 330013, China

<sup>4</sup>School of Statistics, Jiangxi University of Finance and Economics, Nanchang 330013, China

Correspondence should be addressed to Yawei Qi; [qiyawei@jxufe.edu.cn](mailto:qiyawei@jxufe.edu.cn)

Received 21 January 2022; Accepted 9 February 2022; Published 17 March 2022

Academic Editor: Daqing Gong

Copyright © 2022 Mo Xu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper makes a new attempt to identify the effectiveness of innovation factor allocation with a random forest method. This method avoids the evaluation bias of the relative effectiveness caused by the noneffective selection of production frontier in the nonparametric DEA method. It does not refer to other optimal subjects but shifts the focus to the judgment of its own effectiveness. In addition, it also gets rid of the constraints of the model and variables in the parameter SFA method, ensuring the reliability of the measurement results by resampling thousands of times. The data is collected from 30 provinces in China from 2009 to 2018. The findings show the innovation factor allocation in more than half of the provinces is not fully effective. It indicates that how to make use of innovation factor inputs to achieve the actual innovation output higher than own optimal levels is currently still in a period of exploration in China. To further improve innovation factor allocation efficiency, it deeply analyzes the impacts of innovation factor inputs and finds out the important innovation factor inputs. Furthermore, this study presents the nonlinear characteristics and optimal combination of important innovation factor inputs. According to this, it offers the detailed suggestions about how to adjust current important innovation factor inputs for each province in order to greatly enhance the effectiveness of innovation factor allocation in the future.

## 1. Introduction

After the outbreak of the new crown pneumonia epidemic, the global supply chain is forced to terminate, international and domestic demand are blocked, the barriers to the flow of factors are strengthened, and the phenomenon of factor misplacement aggravates. And the five major technologies of big data, cloud computing, Internet of Things, blockchain, and artificial intelligence emerge contribute to the vigorous development of the real economy and bring the great value of the digital economy. As a basic resource in the digital economy, data factors promote the transformation and upgrading of other production factors. The optimization of the spatial layout of regional innovation factors is inseparable from the support and guidance of government policies. This not only reflects that new innovation factors such as

data and institution have significantly affected the transformation of productivity and production relations, but also broadens the connotation scope of the original traditional innovation factors. In addition, the distribution of innovation factor inputs is uneven and innovation capability is generally low among provinces in China, while the calculation results of innovation efficiency are controversy. It is of great practical significance to identify the effectiveness of innovation factor allocation in each province under the new connotation of innovation factors. To this end, clarifying the multidimensional characteristics of innovation factors, selecting reliable measurement methods, and analyzing the current situation of innovation factor allocation in each province play a powerful role in stimulating the innovation factors potential and driving economic growth.



The existing literature measures the allocation efficiency of innovation factors roughly from the following three levels: First, different innovation inputs and innovation outputs. Duan et al. [1] construct a comprehensive index of innovation development from the three dimensions of innovation inputs, innovation organization, and innovation outputs. Dai et al. [2] use financial resources as scientific and technological financial inputs, and scientific and technological achievements as scientific and technological financial outputs. Tao and Peng [3] regard internal R&D expenditures and full-time equivalent of R&D personnel as technological innovation inputs, and the number of patent applications received, revenue from new product sales, and revenue from main operations as technological innovation outputs. Second, different innovation efficiency evaluation objects and various innovation efficiency measurement ways. Li et al. [4] decompose the innovation process into innovation development process and innovation transformation process and then calculate the innovation efficiency of 37 subsectors in China from 2004 to 2011. Wang and Lan [5] use the SFA model to estimate the innovation efficiency of China's A-share listed state-owned enterprises. Chen and Li [6] adopt a two-stage network DEA method to obtain the green innovation efficiency, green technology research and development efficiency, and green technology achievement conversion efficiency in China from 2003 to 2017. Third, measure the allocation efficiency by the factor allocation distortion. Dollar and Wei [7] find that the systematic distortion of capital allocation in China leads to uneven marginal returns across companies, regions, and departments. They point out that if capital can be allocated more effectively, it can lower capital intensity by 5% without sacrificing economic growth. Hsieh and Klenow [8] find that if capital and labor are redistributed so that the marginal output of China and India is equal to that of the United States, this will increase the total factor productivity of the manufacturing industries of the two countries, respectively, by 30%–50% and 40%–60%. Chen and Hu [9] show that resource misallocation leads to a gap of about 15% between actual output and optimal output in various subsectors of China's manufacturing industry.

Existing researches which discuss the allocation modes of innovative factors mainly can be divided into internal, external, and open innovation. Internal innovation means that the innovation subject relies on its own innovation factors to achieve innovative production. External innovation means that the innovation subject completes innovation activities with the help of external innovation factors. Open innovation means that the innovation subject breaks through its own innovation boundaries and coordinates internal and external innovation factors in the innovation process [10]. However, different ways of allocating innovation factors have different impacts on innovation performance. Hong and Shi [11] believe that independent R&D is significantly positively correlated with enterprise innovation performance. Chen and Hou [12] find that the influence of independent innovation on scientific and technological performance manifests a nonlinear threshold effect. Kafourous et al. [13] conclude that the impact of

academic cooperation on the innovation performance in Chinese emerging market companies is significantly positive. Asimakopoulou et al. [14] point out that there is an inverted U-shaped relationship between external knowledge and enterprise innovation efficiency. Berchicci (2013) [10] believes that the inverted U-shaped relationship between external R&D activities and enterprise innovation performance is the function of the substitution effect of its own R&D capabilities. Nieves and Petra [15] find that internal and external knowledge have substitution effects within a certain threshold and have a negative impact on enterprise innovation performance. When the threshold is exceeded, it will turn into a complementary effect and has a positive impact on the enterprise innovation performance.

In summary, the definition of the connotation of innovation factors in the existing literature is mostly based on the traditional innovation factors such as labor, capital, and technology. But it lacks the content of new innovation factors such as data and institution. The existing literature ignores the theoretical basis and operation process behind the innovation factor allocation. The efficiency measurement of the innovation factor allocation in the existing literature is mostly by means of the parametric SFA and the nonparametric DEA. In addition, there are few studies on the evaluation of innovation factor allocation effectiveness. Adopting a more reliable method to identify the effectiveness of innovation factor allocation is helpful to evaluate the situation of innovation factor allocation more accurately. In view of this, the main contributions of this article are as follows: On the one hand, with the introduction of the data and institution innovation factors, we build the index of multidimensional innovation factor, which not only weighs the internal and external coordination of traditional innovation factor from the cost theory, but also considers the moderating effect of new innovation factor from the coupling theory. Based on this, a four-dimensional integrated indicator is formed including internal innovation factors, external innovation factors, institutional innovation factors, and data innovation factors and the allocation of innovation factor is evaluated on the input-output process of innovation factor. On the other hand, we make a new attempt to identify the effectiveness of innovation factor allocation by a random forest method that is different from the previously used. This method can not only get rid of constraints including the model setting form, the number of variables, and the correlation between variables in the parameter SFA method, but also avoid the evaluation bias of the relative effectiveness caused by the noneffective selection of production frontier in the nonparametric DEA method. It shifts the focus on its own efficiency, ensuring the reliability of the measurement by resampling thousands of times. The above aspects are the contribution of the theoretical level. Furthermore, in the aspect of practical level, we draw many different conclusions from existing studies, which find that the innovation factor allocation in more than half of the provinces is not fully effective, data innovation factor inputs especially data integration and data applications have the significant contribution to innovation output, the marginal impact characteristics of important innovation factor inputs are all

nonlinear, and the gap between the current situation and the optimal combination of important innovation factor inputs in each province in China is obvious, etc.

The remaining structure of this paper is as follows: In Section 2, we present the definition and measurement method of innovation factor allocation. In Section 3, we identify the effectiveness of innovation factor allocation in each province in China and discuss the importance of innovation factor inputs. In Section 4, we determine the important innovation factor inputs, describe their nonlinear characteristics and optimal combination, and then point out the future adjustment direction of innovation factor inputs in each province in China. In Section 5, we offer the main conclusions and future research.

## 2. Innovative Factors: Definition, Indicator, and Measurement Method

### 2.1. Definition

*2.1.1. The Definition of Innovation Factors.* Production factors are the objective basis of production activities, and factors affecting production activities all belong to the category of factors [16]. The narrow defining of production factors can be understood as the factors joining in the production process, while the broad defining of production factors should also include factors that reflect the output benefits of the production process. Schumpeter [17] defines innovation as introducing a new combination of production factors and production conditions into the production system that has never been before. Combining modern production conditions with organizational models, innovation can be understood as not only the realization of new products (services) or procedures, the improvement of original products (services) or procedures, but also the process of new products (services) or procedures that are about to or have been commercialized [18–20].

To sum up, this article defines innovation factors as production factors that participate in the innovation process, influence innovation performance, and reflect innovation achievements. They have the characteristics of increasing marginal returns, which include not only traditional innovation factors such as labor, capital, and technology, but also new innovation factors such as data and institution.

*2.1.2. The Definition of the Multidimensional Innovation Factor Allocation.* The cost theory holds that the criterion for organizing economic activities is the minimization of internal production costs and external transaction costs. Drawing on Nieves and Petra (2014) [15], the innovation factors are divided into internal innovation factors and external innovation factors, the internal innovation factors are derived from the innovation subject itself, the external innovation factors are derived outside the innovation subject, and the internal and external innovation factors are mainly based on traditional innovation factors. On the one hand, the innovation subject relies on internal innovation human, material and financial resources, regards innovation

as an internal control process, completes all innovation links and leaves innovation output within the subject. On the other hand, the innovation subject combines their own existing innovation foundations and regards innovation as a R&D outsourcing process. By contacting external innovation partners, they achieve technology cooperation, absorption, and acquisition of external innovation and enlarge enterprise innovation outputs. If the innovation subject is overconcentrated on internal innovation, it will cause problems such as excessive risk and knowledge spillover, while overreliance on external source innovation will bring about problems such as rising negotiation costs and loss of initiative. Therefore, the innovation subject needs to effectively coordinate internal and external innovation factors to minimize the total cost of innovation activities.

The coupling theory states integrating the coordination channels of internal and external traditional innovation factors, enhancing the coupling viscosity of internal and external traditional innovation factors, and reducing the misallocation of internal and external traditional innovation factors, which require both the participation of innovation subjects themselves and innovation assistance from forces other than the subjects. We refer to analysis of Berchicci [10]. On the one hand, the government-led institutional innovation factors regulate the depth of internal and external innovation factors coordination, integration, and fit by intellectual property protection, financial education support, free trade in goods, market transaction quotas, financial development scale, and industrial pollution control. On the other hand, data innovation factors led by intermediaries regulate the breadth of internal and external innovation factors synergy, integration and adaptation by data production level, data transmission speed, data user groups, data application scope, data sharing degree, and data integration capabilities. Institution and data innovation factors play the regulatory role of institutions and information (intrinsic nature of data) on internal and external innovation factors, thereby forming a multidimensional integrated indicator of internal innovation factors, external innovation factors, institutional innovation factors, and data innovation factors.

Based on the above analysis, this article draws the definition of the multidimensional innovation factor allocation. The cost theory synergizes internal and external innovation factors. The measurement indicator of internal innovation factors can be from the three aspects of internal innovation human resources, material resources, and financial resources. The measurement indicator of external innovation factors can be from three aspects such as external innovation technologies cooperation, absorption, and acquisition. The regulation role of institution and data innovation factors depends on coupling theory. Institution innovation factor measurement indicator can be from six aspects including property rights protection, education support, trade freedom, trading markets, financial development, and pollution control. Data innovation factor measurement indicator can be from six aspects including data generation, transmission, use, application, sharing, and integration.

## 2.2. Indicator

**2.2.1. Indicator Composition.** This paper is based on the definition of innovation factors allocation, the operability of indicators, and the availability of data. Combining with the research of Tao and Xu [21], we select the specific quantitative indicator. The internal innovation human resources, financial resources, and material resources in the internal innovation factors are, respectively, measured from the full-time equivalent of R&D personnel, internal R&D expenditures, and the number of R&D organizations. The external innovation technology cooperation, absorption, and acquisition in external innovation factors are, respectively, measured from external R&D expenditures, foreign direct investment, and high-tech product imports. Property rights protection, education support, trade freedom, trading market, financial development, and pollution control in the institution innovation factors are, respectively, measured from the total number of patent enforcement cases, financial education subsidies, total regional goods exports, technology market turnover, deposits and loans of financial institutions, and industrial pollution control investment. The data generation, transmission, use, application, sharing, and integration in the data innovation factors are, respectively, measured from the number of Internet pages, Internet broadband access ports, Internet users, total software services, total post and telecommunications services, and the number of robots owned. Taking the above four dimensions of internal, external, institution, and data innovation factors as innovation inputs, and taking per capita sales income of new products as innovation output, we evaluate China's innovation factors allocation efficiency.

**2.2.2. Variables and Data.** This paper is based on the panel data of 30 provinces in China from 2009 to 2018 (except Tibet). The data are extracted from China Statistical Yearbook on Science and Technology, China Statistical Yearbook, China Internet Development Report, China Information Yearbook, and International Federation of Robotics. The acquisition of the variables and the processing of the data are performed as follows: Regarding the matches and fillings of the data, the first is the number of robots owned in China based on the industrial output value of the main applications of industrial robots in each province. The second is filling the missing data of provinces according to the difference between the national total and existing provinces total. In addition, with regard to the methods of the processing of the data, firstly, exchange rate adjustments are converted by the ratio of RMB to USD. Secondly, the price adjustment refers to the price index formula of Zhu and Xu [22]. The fixed asset investment index and the consumer price index are, respectively, assigned weights of 0.45 and 0.55. Regarding the year 2009 as the price base period, convert the nominal values into actual values. Thirdly, using the method of BEA for stock adjustment, based on the annual investment flow after price adjustments  $E_t$ , we calculate the average annual growth rate  $g_k$ . Then, considering amortization of intangible assets for at least 10 years, and referring to Wang and Gao [23], set the

residual value rate  $d_T$  to 10% and get the depreciation rate  $\delta = 1 - (d_T)^{1/T} = 0.2057$ . Lastly, the annual capital stock is based on  $K_0 = [E_1(1 - \delta/2)]/[g_k + \delta]$  and  $K_t = (1 - \delta)K_{t-1} + (1 - \delta/2)E_t$ . The above contents are detailed in Table 1.

**2.3. A Random Forest Method.** The existing researches of measuring efficiency mainly use nonparameter DEA method and parameter SFA method. The DEA method is used to evaluate relative efficiency by constructing the optimal production frontier. On the one hand, it refers to other effective subjects instead of its own optimum, which is difficult to purposely offer the adjustment direction according to itself current inputs. On the other hand, it may occur efficiency measurement bias if the optimal production frontier is selected mistakenly. In addition, the SFA method measures efficiency by determining the form of model and the number of variables in advance. It means the differences of the form of model and the number of variables lead to the different measurement results. Meanwhile, the collinearity between variables in the model also causes biased efficiency measurement. Compared with DEA and SFA, the random forest method not only gets rid of constraints including the form of model, the number of variables, and the correlation between variables, but also ensures the reliability of measurement by resampling thousands of times. With reference to Ouyang and Chen [24], the prediction of output can be regarded as the maximum output achieved under the full utilization of inputs by the random forest method. According to this, we can calculate own optimal output in terms of current inputs and then identify effectiveness by the ration of actual output to optimal output. It does not refer to other optimal subjects but shifts the focus to the judgment of its own effectiveness. Moreover, it gets the contribution rankings of various inputs to outputs from large to small based on the principle of minimum sum of squares of residuals. We further get the marginal impact characteristics of inputs and find out the optimal combination of inputs and thus put forward the further adjustment direction depending on own current situation of inputs. The introduction of random forest method is below in detail.

A random forest method is an ensemble learning method. This method uses resampling to randomly select sample data and sample characteristics to establish multiple regression numbers, and the mean value of the output results of multiple regression trees is used as the final prediction result. Based on the data of each input, this paper predicts the results of each output according to the random forest method. The detailed steps are as follows:

- (i) Suppose there are samples  $\{y_i, c_{i1}, \dots, c_{iK}\}_{i=1}^N$ ,  $N$  and  $K$ , respectively, represent total number of samples and features,  $y_i$  is the individual  $i$  output, and  $c_{i1}, \dots, c_{iK}$  is the individual  $i$  input. Based on this, samples are drawn randomly with replacement.
- (ii) Construct a regression tree with features  $m$  selected from the sample, and perform feature splitting according to the principle of the minimum sum of

TABLE 1: Multidimensional innovation factor allocation indicator.

Dimension	Target layer	Operation layer	Data sources	Processing method	
Internal innovation factor	Internal innovation human resources	The full-time equivalent of R&D personnel	China Statistical Yearbook on Science and Technology	—	
	Internal innovation financial resources	Internal R&D expenditures	China Statistical Yearbook on Science and Technology	Price adjustment and stock adjustment	
	Internal innovation material resources	The number of R&D organizations	China Statistical Yearbook on Science and Technology	—	
External innovation factors	External innovation technology cooperation	External R&D expenditures	China Statistical Yearbook on Science and Technology	Price adjustment and stock adjustment	
	External innovation technology absorption	Foreign direct investment	China Statistical Yearbook	Exchange rate adjustment, price adjustment, and stock adjustment	
	External innovation technology acquisition	High-tech product imports	China Statistical Yearbook on Science and Technology	Exchange rate adjustment and price adjustment	
	Property rights protection	The total number of patent enforcement cases	China Statistical Yearbook	—	
Inputs	Education support	Financial education subsidies	China Statistical Yearbook	Price adjustment	
	Institution innovation factors	Trade freedom	Total regional goods exports	China Statistical Yearbook	Exchange rate adjustment and price adjustment
		Trading market	Technology market turnover	China Statistical Yearbook	Price adjustment
		Financial development	Deposits and loans of financial institutions	China Statistical Yearbook	Price adjustment
	Pollution control	Industrial pollution control investment	China Statistical Yearbook	Price adjustment	
	Data innovation factors	Data generation	The number of Internet pages	China Statistical Yearbook	—
		Data transmission	Internet broadband access ports	China Statistical Yearbook	—
		Data use	Internet users	China Internet Development Report and China Statistical Yearbook	—
		Data application	Total software services	China Information Yearbook	Price adjustment
		Data sharing	Total post and telecommunications services	China Statistical Yearbook	Price adjustment
Data integration	The number of robots owned	International Federation of Robotics	Price adjustment		
Output	Innovation output	Per capita sales income of new products	China Statistical Yearbook on Science and Technology	Price adjustment	

Note: all the data are processed logarithmically to eliminate the influence of differences in indicator units.

squared residuals. Specifically for the input indicator  $c_k$ , the critical point at which the optimal threshold is known is  $d$  (the determination of  $d$  also adopts the principle of minimum sum of squares of residuals). Then according to the sample size of the left and right sides of the critical points  $T_{U1}$  and  $T_{U2}$ , we, respectively, get the sample of  $U1$  and  $U2$ :

$$\begin{aligned} U1(k, d) &= \{y_i, c_{i1}, \dots, c_{iK}/c_k \leq d\}_{i=1}^{T_{U1}}, \\ U2(k, d) &= \{y_i, c_{i1}, \dots, c_{iK}/c_k > d\}_{i=1}^{T_{U2}}. \end{aligned} \quad (1)$$

From this, the average value of the output results of each sample can be obtained:

$$\bar{y}_{U1} = \text{mean}(y_i/y_i \in U1), \quad \bar{y}_{U2} = \text{mean}(y_i/y_i \in U2). \quad (2)$$

The selection of the initial split node  $c_k$  of the regression tree is based on the following formula:

$$\min_{C_k} \left[ \min \sum_{i=1}^{T_{U1}} (y_i - \bar{y}_{U1})^2 + \min \sum_{i=1}^{T_{U2}} (y_i - \bar{y}_{U2})^2 \right]. \quad (3)$$

- (i) The selection of the remaining nodes of the regression tree is repeated based on step (ii). The prepruning rule is set to include at least 5 sample points for each node. When the rules are met, the regression tree immediately stops splitting.

- (ii) The loop steps (i)-(iii) form a large number of regression trees, and the average value of the above regression tree results is used as the final prediction of the output results.

In addition to regression analysis, the random forest method can also evaluate the importance of each input and its marginal influence on the output. The nodes of the regression tree are arranged from top to bottom according to the contribution of each input to the reduction of the residual sum of squares of the output. The top is the input with the largest contribution, and the bottom is the input with the least contribution. According to this, the importance of each input can be obtained. By sorting the importance of input, we can get the important order of input. The marginal effect is mainly used to measure the impact of a single input change on the output, which can be defined as

$$\bar{f}(c_k) = \frac{1}{N} \sum_{i=1}^N f(c_{i1}, \dots, c_k, \dots, c_{iK}). \quad (4)$$

In equation (4),  $f(c_{i1}, \dots, c_k, \dots, c_{iK})$  is the prediction result of the random forest model  $f(\cdot)$  when  $\{c_{i1}, \dots, c_k, \dots, c_{iK}\}_{i=1}^N$  is known.  $\bar{f}(c_k)$  is the mean of the predicted results for all samples. Traverse all possible values  $c_k$  to get the corresponding result  $\bar{f}(c_k)$ . This draws a trend graph  $\bar{f}(c_k)$  about  $c_k$  and visualizes the influence path of the input  $c_k$  on the mean value of output  $\bar{f}(c_k)$ .

### 3. Innovation Factors: Allocation Effectiveness and Importance Degree

*3.1. Effectiveness Identification.* We use random forest method, the algorithm is run 1000 times, and the average value of 1000 times is regarded as the final prediction of innovation output. Based on the above analysis, it is also the optimal innovation output achieved by innovation factor inputs. Then we get the ration of actual output to optimal output, which is the allocation efficiency of innovation factors from 30 provinces in China from 2009 to 2018. Based on this, we identify the effectiveness of innovation factor allocation for each province in China and offer the corresponding schematic graph shown in Figure 1.

In Figure 1, the black solid dots indicate that the province's actual innovation output exceeds its own optimal output, realizing the effective allocation of innovation factors, while the gray solid dots indicate that the province's actual innovation output is lower than its own optimal output. The effective allocation of innovation factors has not been completed. According to the allocation of innovation factors in each province from 2009 to 2018, each province is divided into the following three types: one is fully effective, and the province achieves an effective allocation of innovation factors during the entire sample period; the second is not fully effective, and the province only achieves the effective allocation of innovation factors in a part of the sample period; the third is totally ineffective, and the province does not meet the requirements of the effective allocation of innovation factors during the full sample period.

As can be seen from Figure 1, the provinces with innovation factor allocation fully effective are Tianjin, Shanxi, Shanghai, Zhejiang, Hubei, Hunan, Chongqing, and Ningxia. The above provinces can not only effectively coordinate internal and external innovation factors during the entire sample period, but also give full play to the coupling role of institution and data innovation factors, organically integrate the four-dimensional innovation factors into one, and obtain innovation output that exceeds its own best state. Although Chongqing and Ningxia are located in the relatively backward western regions, there is a certain gap in the innovation factor inputs compared with the developed regions, and the optimal output that can be achieved by the current innovation factor inputs is relatively low, so the actual output of the two provinces reaching the optimal level is also relatively easy. The provinces where the allocation of innovation factors is not fully effective are Beijing, Inner Mongolia, Liaoning, Jilin, Jiangsu, Anhui, Fujian, Jiangxi, Shandong, Guangdong, Guangxi, Gansu, Qinghai, and Xinjiang. In the abovementioned provinces, the innovation output obtained by the innovation factor inputs in only some years is higher than its own optimal level. Although Beijing, Jiangsu, Fujian, Shandong, and Guangdong are located in eastern regions with superior innovation conditions and abundant innovation resources, their innovation factor inputs are relatively more, and the optimal output that can be achieved by relying on their existing innovation factor inputs is also relatively high. The actual output beyond its own optimal level is also relatively difficult. The provinces where the allocation of innovation factors is totally ineffective are Hebei, Heilongjiang, Henan, Hainan, Sichuan, Guizhou, Yunnan, and Shaanxi. The actual output of the above eight provinces during the entire sample period was lower than their optimal level. Although Hebei is located in the economically developed eastern region, the relatively high innovation factor inputs have caused structural imbalance, and it is difficult to fully release the efficacy of innovation factor inputs.

*3.2. Importance Judgment.* From the foregoing, the random forest method can get the importance of the innovation factor inputs according to the contribution to the reduction of the square of the residual of the innovation output. It puts the most important innovation factor input at the top of the tree and the least important innovation factor input at the bottom of the tree. Therefore, the higher the innovation factor input located at the level of the tree, the greater the contribution to the innovation output. According to the importance of the innovation factor inputs from 2009 to 2018 (see Table 2), this paper draws a grid graph and a boxplot graph (see Figures 2 and 3). The grid graph reflects the contribution degree, the changing trend, and the abnormal time point. The larger the grid area, the greater the importance. The boxplot presents the average contribution and scatter degree. Based on the above analysis, it provides clear direction for improving the effectiveness of innovation factor allocation.

In Table 2 and Figure 2, it can be seen that the contribution of data application is the most prominent in 2009

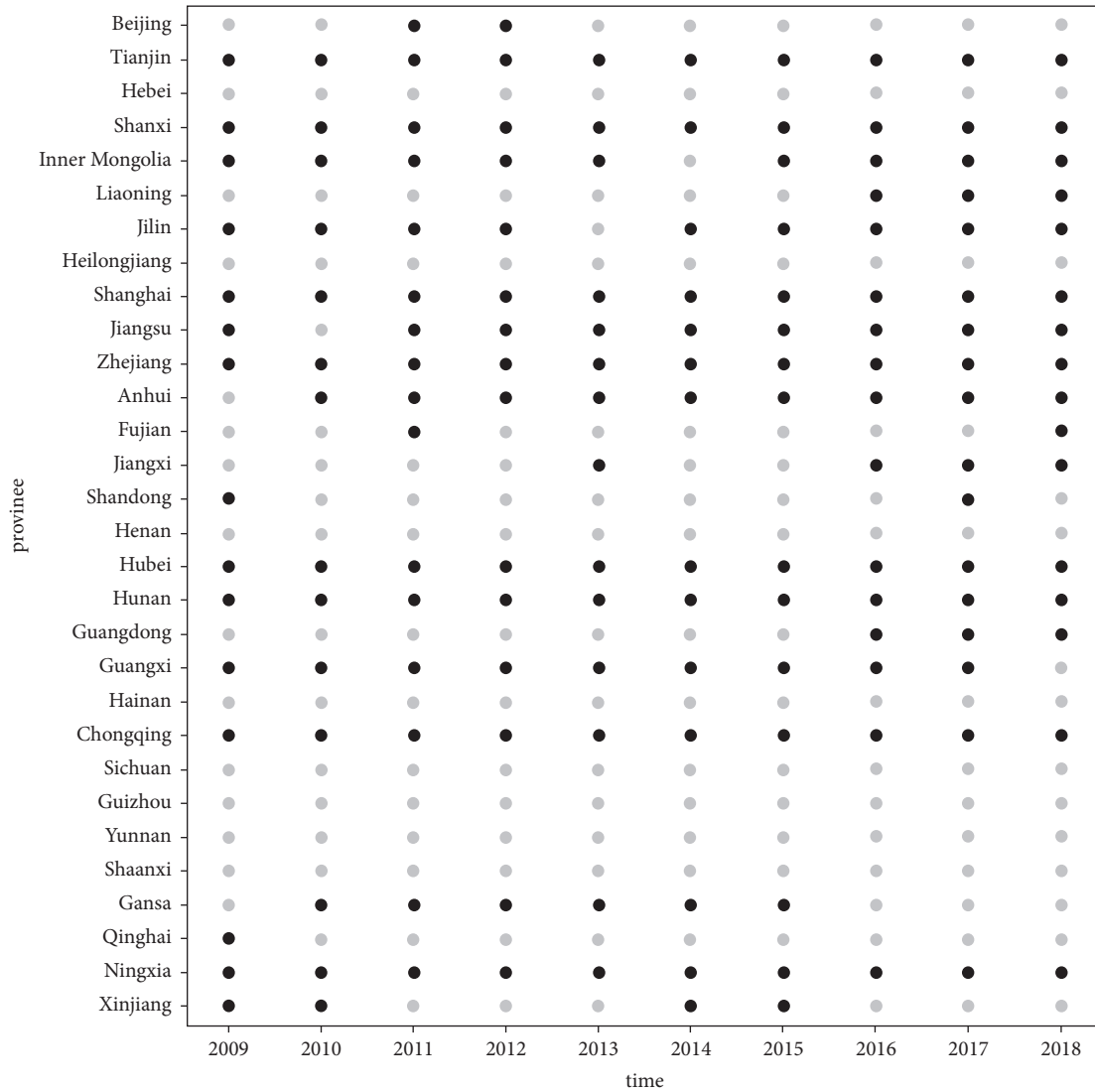


FIGURE 1: Effectiveness of innovation factor allocation.

TABLE 2: Importance of innovation factor inputs.

Innovation factor inputs	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Mean	Ranking
Internal innovation human resources	0.0698	0.0543	0.0693	0.0742	0.0669	0.0579	0.0678	0.0706	0.0639	0.0961	0.0691	7
Internal innovation financial resources	0.0519	0.0530	0.0471	0.0523	0.0368	0.0567	0.0505	0.0600	0.0575	0.0588	0.0525	9
Internal innovation material resources	0.0878	0.0704	0.0636	0.0710	0.0831	0.0894	0.1036	0.1174	0.1105	0.0697	0.0867	4
External innovation technology cooperation	0.0384	0.0455	0.0463	0.0409	0.0502	0.0511	0.0518	0.0453	0.0376	0.0560	0.0463	11
External innovation technology absorption	0.0636	0.0873	0.0833	0.0912	0.1010	0.0824	0.0917	0.0928	0.1064	0.1367	0.0936	3
External innovation technology acquisition	0.0627	0.1031	0.0888	0.0842	0.0742	0.0786	0.0583	0.0616	0.0792	0.0723	0.0763	5
Property rights protection	0.0314	0.0124	0.0173	0.0275	0.0228	0.0235	0.0205	0.0163	0.0158	0.0197	0.0207	16
Education support	0.0183	0.0187	0.0127	0.0132	0.0145	0.0143	0.0146	0.0130	0.0142	0.0149	0.0148	18
Trade freedom	0.0625	0.0560	0.0598	0.0754	0.0766	0.0595	0.0625	0.0682	0.0547	0.0661	0.0641	8
Trading market	0.0692	0.0926	0.0613	0.0241	0.0256	0.0216	0.0198	0.0292	0.0270	0.0294	0.0400	12



TABLE 2: Continued.

Innovation factor inputs	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Mean	Ranking
Financial development	0.0551	0.0540	0.0623	0.0657	0.0957	0.0837	0.0717	0.0623	0.0787	0.1123	0.0742	6
Pollution control	0.0188	0.0112	0.0156	0.0195	0.0204	0.0094	0.0153	0.0158	0.0267	0.0275	0.0180	17
Data generation	0.0439	0.0391	0.0569	0.0491	0.0627	0.0589	0.0680	0.0386	0.0305	0.0378	0.0485	10
Data transmission	0.0592	0.0338	0.0367	0.0320	0.0439	0.0367	0.0334	0.0227	0.0320	0.0217	0.0352	13
Data use	0.0287	0.0190	0.0201	0.0278	0.0276	0.0250	0.0192	0.0175	0.0182	0.0165	0.0220	15
Data application	0.1078	0.1226	0.1138	0.1008	0.0707	0.0933	0.0876	0.0924	0.0883	0.0653	0.0943	2
Data sharing	0.0294	0.0245	0.0301	0.0301	0.0383	0.0350	0.0324	0.0374	0.0223	0.0215	0.0301	14
Data integration	0.1015	0.1025	0.1150	0.1211	0.0889	0.1227	0.1313	0.1387	0.1367	0.0777	0.1136	1

Note: the ranking results are based on the annual average value.

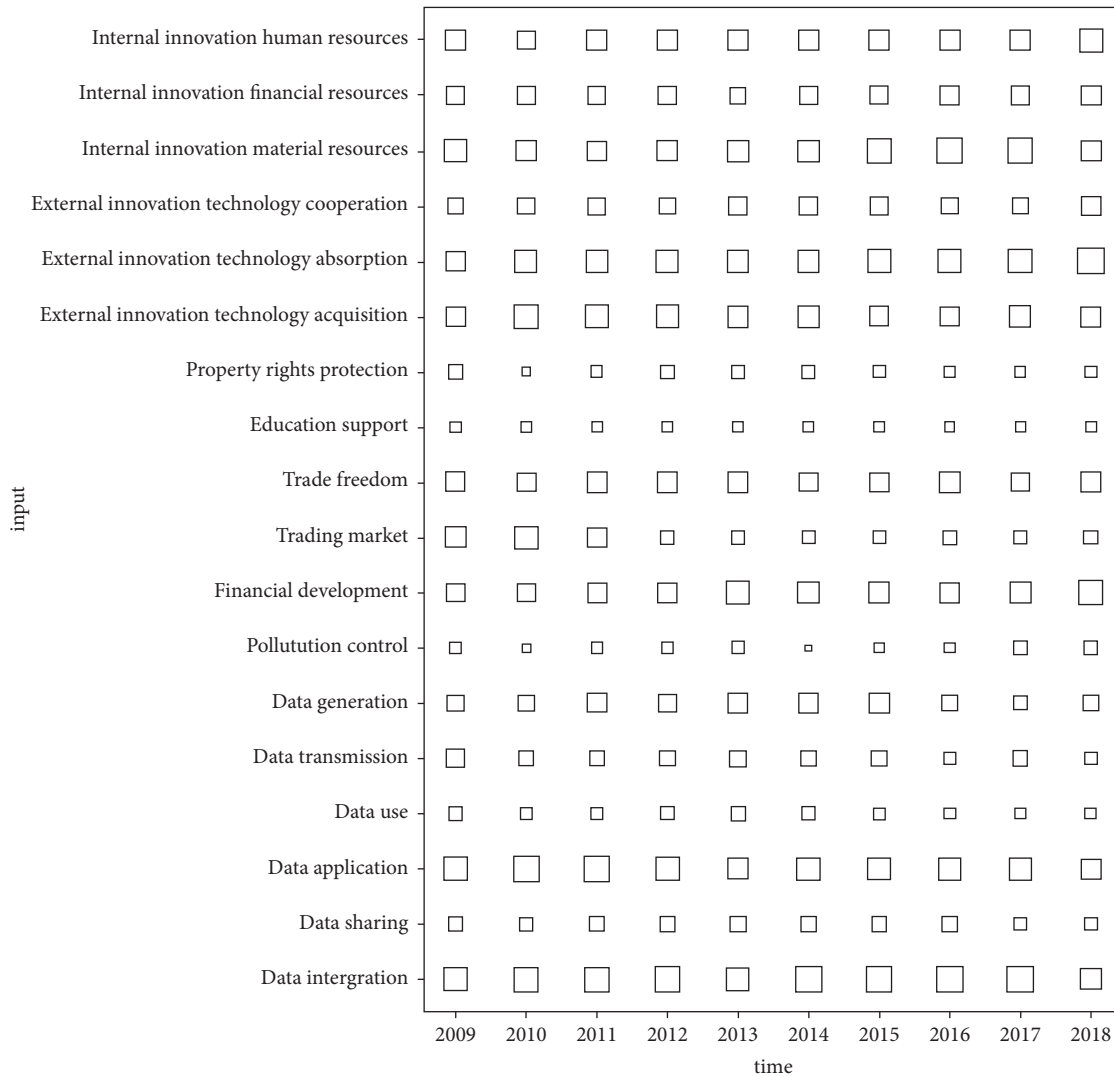


FIGURE 2: Grid graph of innovation factor inputs importance.

and 2010, the contribution of external innovation technology absorption is the most significant in 2013 and 2018, and the performance of data integration is the most eye-catching in the remaining years, which show that the data application, external innovation technologies absorption, and data integration have significant impacts on the innovation output during the entire sample period. In addition, the number

when the input of innovation factors contributes the most value itself is as follows. First, the number topping the list in 2018 is five, respectively, including internal innovation human resources, external innovation technology cooperation, external innovation technology absorption, financial development, and pollution control. Second, the number in 2010 is four, respectively, including external innovation

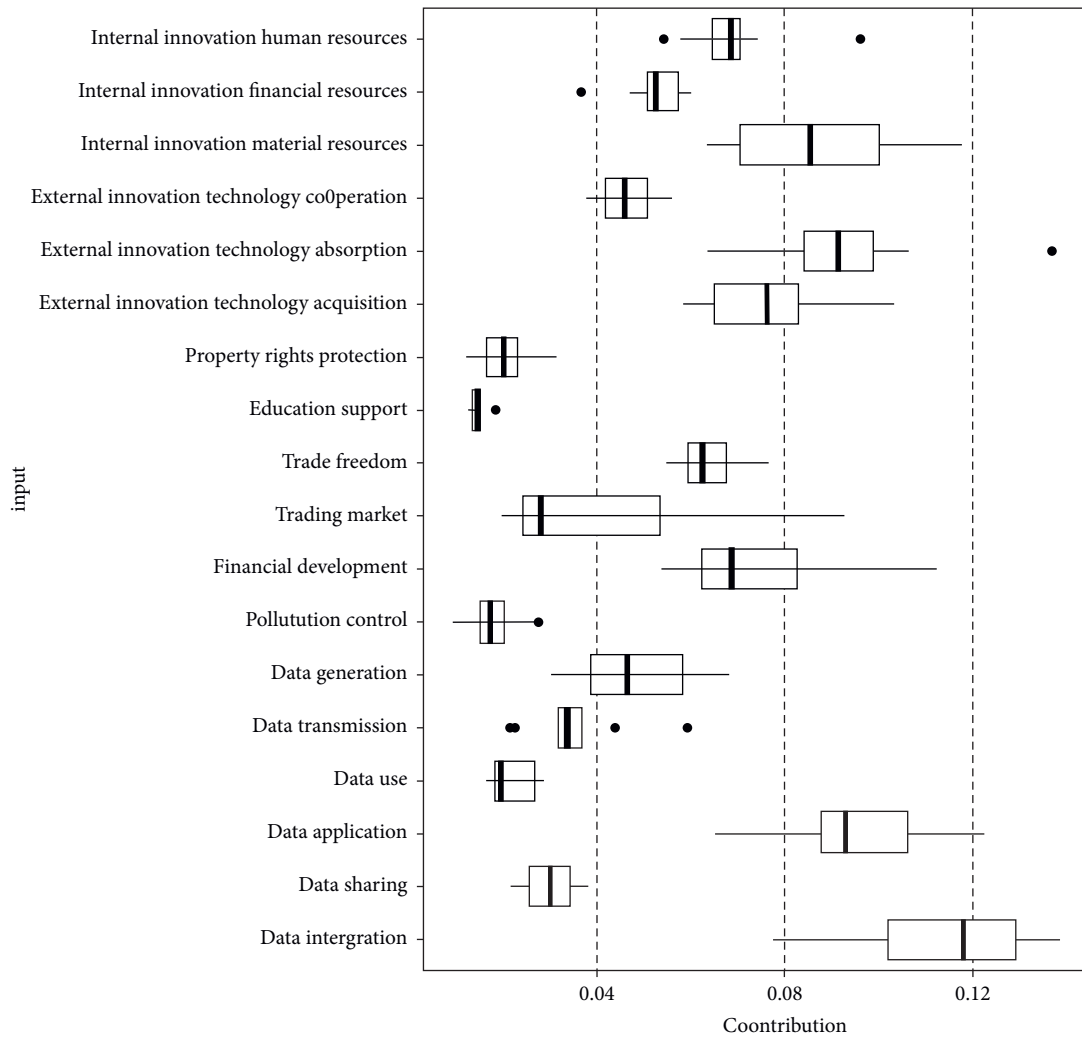


FIGURE 3: Boxplot of innovation factor inputs importance.

technology acquisition, education support, trading market, and data application. Third, the number in 2009 and 2016 is three, which, respectively, include property rights protection, data transmission and data use in 2009, and internal innovation financial resources, internal innovation material resources, and data integration in 2016. Forth, the number in 2013 is two, respectively, including trade freedom and data sharing, the input in 2015 is data generation, and there is none in 2011, 2012, 2014, and 2017. This shows that the input structure of innovation factors in 2018 is more reasonable than other years, it gives full play to the advantages of innovation factor inputs, and there is the most innovation factor input contributing its greatest impact.

In Table 2 and Figure 3, the average contribution of innovation factor inputs during the entire sample period is ranked from large to small, followed by data integration, data application, external innovation technology absorption, internal innovation material resources, external innovation technology acquisition, financial development, internal innovation human resources, trade freedom, internal innovation financial resources, data generation, external innovation technology cooperation, trading market, data transmission,

data sharing, data use, property rights protection, pollution control, and education support. The ranking top two belong to data innovation factors. It confirms not only the importance of the data innovation factor inputs, but also the data integration measured by the number of robots owned and the data application measured by the total amount of software business, which implies artificial intelligence and advanced manufacturing playing the important role within reforming the synergy of internal and external innovation factors and promoting the acceleration of innovation output. Data innovation factors have the characteristic of strong integration. They can overcome the limitations of time and space, which is conducive to the optimization and upgrading of internal and external innovation factors and the complementary advantages between internal and external innovation factors, which greatly promotes the improvement of innovation output. Internal innovation human resources, internal innovation financial resources, external innovation technology absorption, education support, pollution control, and data transmission all have outliers. Five have large outliers which all appear at the largest time point of its own contribution, and three appeared in 2018. It not only illustrates the relatively

reasonable input structure of innovation factors in 2018, but also implies that the overall input structure of innovation factors in China is undergoing continuous adjustment and improvement.

#### 4. Innovation Factors: Marginal Effects and Input Choices

*4.1. Marginal Effect Trend.* The above discusses the importance of each innovation factor input to the innovation output under the existing innovation factor inputs combination, which is the static analysis of the impact on the innovation output. If the innovation factor input has a significant impact on innovation output, will it continue to increase inputs or be subject to threshold constraints? The answer to this question requires us to give the dynamic changes in the marginal effect trend of innovation factor input.

Based on the importance ranking of innovation factor inputs, the top eight cumulative impacts exceeding 60% of the total impact are regarded as important innovation factor inputs, followed by data integration, data application, external innovation technology absorption, internal innovation material resources, external innovation technology acquisition, financial development, internal innovation human resources, and trade freedom. At the same time, it is found that the above inputs are evenly distributed in the internal, external, institution, and data innovation factors. In order to obtain the time changing trend of the marginal effects of each important innovation factor, the marginal effect graph of each important innovation factor input is given in the order of the abovementioned four-dimensional innovation factors (see Figure 4). We select 2009 (dotted line type), 2012 (lower solid line), 2015 (dashed line), and 2018 (higher solid line) for describing in detail.

There are three key boundary points in the marginal effect trend graph in Figure 4, which are two turning points (indicated by  $\times$ ) and the maximum slope point (indicated by  $*$ ). Before the first turning point is the initial stage of innovation factor input impact, and the marginal impact on innovation output is relatively flat. The first turning point to the maximum slope point is the marginal impact on innovation output from the increasing to the maximum. The stage between the maximum slope point and the second turning point is the marginal incremental impact of the innovation factor input fallen to the suspended. The second turning point is the innovation factor input required for maximum innovation output. The marginal impact of increasing the innovation factor input on innovation output is almost zero after the second turning point. It can be concluded that the optimal innovation factor is the amount of innovation factor input required when the innovation output reaches its maximum output. If all the important innovation factor inputs in a province have exceeded the second turning point, then the province's important innovation factor investment portfolio is the best. Based on this, we give the key boundary points of the marginal effect of

important innovation factor inputs in 2009, 2012, 2015, and 2018 (see Table 3).

Combining the time trend of the marginal effects of important innovation factors, we find that the impact on innovation output is nonlinear in whole process. In addition, the positions of the first turning point, the maximum slope point, and the second turning point are not exactly the same in 2009, 2012, 2015, and 2018. It can be seen that the demand for important innovation factor inputs in different periods is somewhat different. Details are as follows: Firstly, the first turning point of internal innovation human resources, financial development, and data integration between 2009 and 2015 is relatively closer. The distance of the first turning point in internal innovation material resources, external innovation technology acquisition, and data application between 2012 and 2015 is the shortest. At the same time, except for trade freedom, the first turning point of the other important innovation factor inputs in 2018 appears later than the other years. It shows that the amount of the important innovation factor inputs required at the starting point of increasing marginal effects in 2009, 2012, and 2015 is relatively stable, while the amount of the important innovation factor inputs required in 2018 is higher than that in 2009, 2012, and 2015. Secondly, the distance of all important factor inputs between the first turning point and the maximum slope point in 2018 is smaller than in 2009, 2012, and 2015. It indicates that, compared with other years, the important innovation factor inputs have the certain foundation and do not have a relatively long period of marginal effect growth in turn in 2018. Thirdly, the second turning point of internal innovation human resources in 2018 appears earlier than in 2009, 2012, and 2015, while the second turning point of the other important innovation factors in 2018 appears later than in 2009, 2012, and 2015. It indicates that, comparing 2018 with the other years, the demand for internal innovation human resources is relatively small, and the demand for other important innovation factors is greater in the optimal innovation factor inputs required to achieve the maximum innovation output.

*4.2. Inputs Selection Analysis.* The ranking top eight of the innovation factor inputs' importance in 2012 and 2018 are consistent with the result of the average contribution ranking in Table 2. Therefore, the input of important innovation factors in each province in 2012 and 2018 is carried out by grouping accordingly. It can not only analyze the time changes of important innovation factor inputs, but also propose future adjustment direction based on the current innovation factor inputs in each province.

According to the grouping of important innovation factor inputs in each province in 2012 and 2018 (see Table 4), combined with the marginal effect graph of important innovation factor inputs in 2012 and 2018 (see Figures 4(a)–4(h), respectively, corresponding to the lower and the higher solid line), the following important features are found: Firstly, part of important innovation factor inputs in some provinces have not exceeded the critical value of the first turning point. That means these important innovation factor

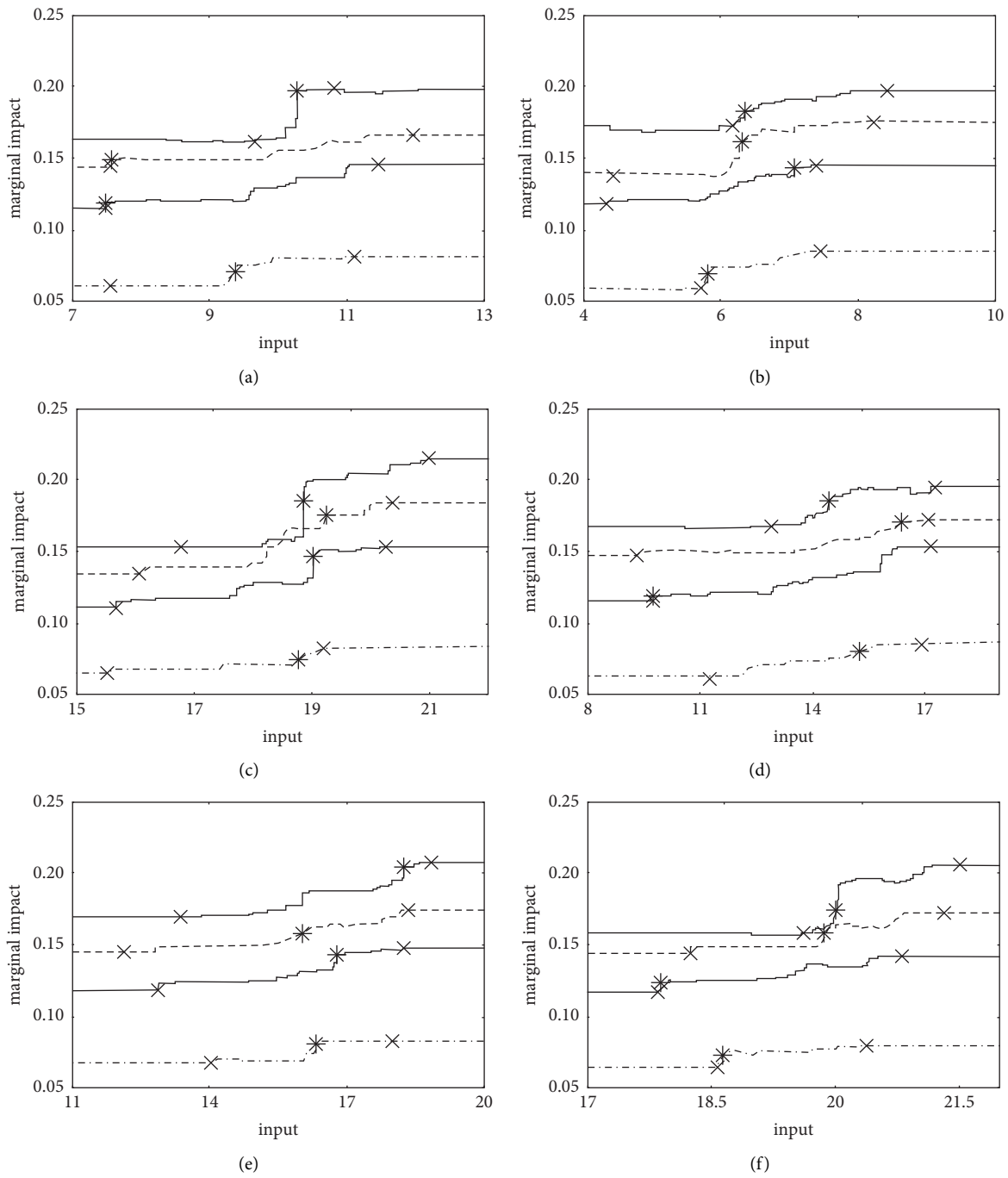


FIGURE 4: Continued.

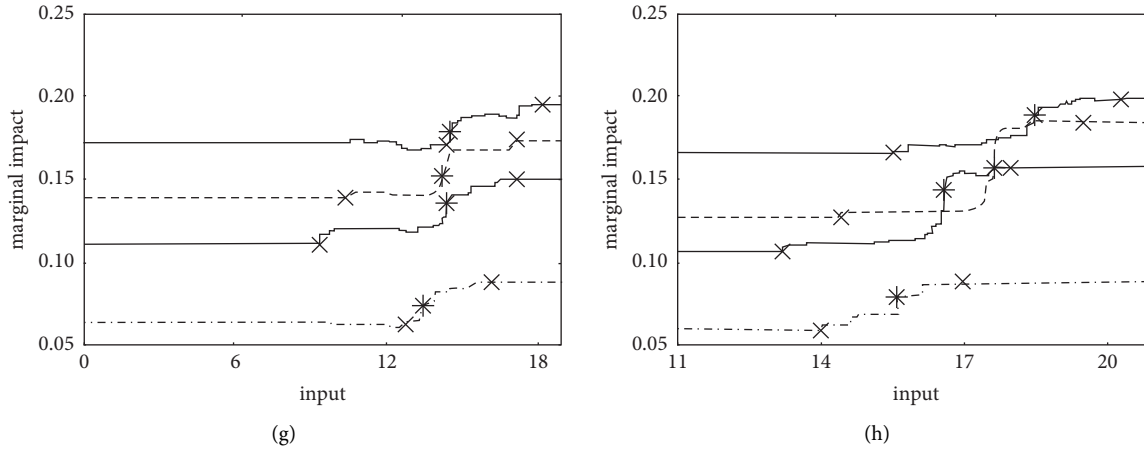


FIGURE 4: Marginal effect trend of important innovation factor inputs. (a) Internal innovation human resources. (b) Internal innovation material resources. (c) External innovation technology absorption. (d) External innovation technology acquisition. (e) Trade freedom. (f) Financial development. (g) Data application. (h) Data integration.

TABLE 3: Key boundary point of the important innovation factor inputs' marginal effect.

	2009			2012			2015			2018		
	$\alpha$	$\beta$	$\gamma$	$\alpha$	$\beta$	$\gamma$	$\alpha$	$\beta$	$\gamma$	$\alpha$	$\beta$	$\gamma$
(a)	7.57	9.39	11.11	7.49	7.50	11.45	7.57	7.58	11.96	9.65	10.28	10.80
(b)	5.70	5.81	7.44	4.34	7.06	7.38	4.44	6.29	8.22	6.17	6.33	8.42
(c)	15.54	18.78	19.20	15.68	19.03	20.24	16.07	19.24	20.38	16.79	18.86	20.97
(d)	11.25	15.21	16.92	9.75	9.76	17.16	9.31	16.35	17.11	12.90	14.42	17.25
(e)	14.02	16.32	17.99	12.88	16.77	18.24	12.15	16.00	18.31	13.36	18.22	18.83
(f)	18.58	18.64	20.38	17.87	17.88	20.80	18.25	19.88	21.31	19.61	20.00	21.51
(g)	12.83	13.51	16.17	9.35	14.36	17.11	10.35	14.19	17.21	14.38	14.54	18.23
(h)	13.99	15.61	16.96	13.20	16.56	17.95	14.42	17.59	19.48	15.50	18.44	20.26

Note:  $\alpha$ ,  $\beta$ ,  $\gamma$ , respectively, represent the first turning point, the maximum slope point, and the second turning point boundary value.

inputs have not yet reached the starting point of increasing marginal effects. Secondly, all important innovation factor inputs in some provinces are between the first turning point and the second turning point. That indicates all important innovation factor inputs are in the stage where their marginal effect increases to the maximum and then to the gradual decline. Thirdly, part of important innovation factor inputs in some provinces have reached the second turning point threshold value, which means that these important innovation factors have completed the optimal inputs. Fourthly, all important innovation factor inputs in some provinces exceed the second turning point threshold value, which indicates that all important innovation factor inputs meet the optimal portfolio requirements. Based on the above four important characteristics, we classify the input of important innovation factors in each province in 2012 and 2018, analyze the characteristics of changes in the important innovation factor inputs, and propose future innovation development policies.

Based on the summary of important innovation factor inputs grouping characteristics in each province in 2012 and 2018 (see Table 5), the changes in the following grouping characteristics of each province are obtained: Firstly, the important innovation factor inputs in some provinces in

2012 and 2018 have always maintained the characteristics (a), such as Qinghai and Ningxia. Although Ningxia shows the innovation factor allocation is effective in 2012 and 2018, the optimal innovation output achieved is low due to its existing innovation factor inputs. It implies that the above provinces need to significantly increase the important innovation factor inputs, which can not only increase the actual innovation output, but also raise the optimal output threshold standard. Secondly, the important innovation factor inputs in some provinces in 2012 and 2018 remain at the characteristic (b), such as Tianjin, Hebei, Liaoning, Jilin, Anhui, Henan, Hunan, and Chongqing. The above provinces may refer to the contribution ranking of important innovation factor inputs and determine the appropriate direction for the increase in the important innovation factor inputs. There are also some provinces where the important innovation factor inputs have dropped from the characteristics (b) in 2012 to the characteristics (a) in 2018, such as Shanxi, Inner Mongolia, Heilongjiang, Jiangxi, Guangxi, Hainan, Guizhou, Yunnan, Gansu, and Xinjiang. The above provinces input less in data applications. The contribution of data application input is among the best. Increasing the inputs of data application can not only make the provinces get rid of the characteristics (a), but also promote the improvement of

TABLE 4: Important innovation factor inputs grouping.

		(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
Beijing	2012	4	3	3	3	2	3	4	3
	2018	4	3	3	3	2	3	3	3
Tianjin	2012	3	2	3	3	3	3	3	3
	2018	3	3	3	3	2	3	3	3
Hebei	2012	3	2	2	3	2	3	2	3
	2018	3	3	3	2	2	3	3	3
Shanxi	2012	3	2	2	3	2	3	2	2
	2018	2	1	2	3	2	3	1	2
Inner Mongolia	2012	3	2	2	3	2	3	2	2
	2018	1	1	2	1	2	2	1	2
Liaoning	2012	3	2	3	3	3	3	3	3
	2018	3	3	3	3	2	3	3	2
Jilin	2012	3	2	2	3	2	3	3	3
	2018	2	2	2	2	2	2	3	3
Heilongjiang	2012	3	2	2	3	2	3	2	2
	2018	2	1	2	1	2	2	1	2
Shanghai	2012	3	3	4	4	4	3	3	3
	2018	4	3	4	4	3	3	3	3
Jiangsu	2012	4	4	4	4	4	4	4	4
	2018	4	4	4	4	4	3	3	4
Zhejiang	2012	3	4	3	3	4	4	3	3
	2018	4	3	3	3	3	3	3	3
Anhui	2012	3	2	2	3	2	3	2	3
	2018	3	3	3	3	2	3	3	3
Fujian	2012	3	2	3	3	3	3	3	3
	2018	4	3	3	3	2	3	3	3
Jiangxi	2012	3	2	2	3	2	3	2	2
	2018	3	3	3	3	2	3	1	3
Shandong	2012	3	4	3	3	3	3	3	4
	2018	4	3	3	3	2	3	3	3
Henan	2012	3	2	2	3	2	3	2	3
	2018	3	3	3	3	2	3	3	3
Hubei	2012	3	2	2	3	2	3	3	3
	2018	4	3	3	3	2	3	3	3
Hunan	2012	3	2	2	3	2	3	3	3
	2018	3	3	3	3	2	3	3	3
Guangdong	2012	4	4	4	4	4	4	4	4
	2018	4	4	4	4	4	4	4	4
Guangxi	2012	3	2	2	3	2	3	2	2
	2018	2	2	2	3	2	2	1	2
Hainan	2012	3	2	2	3	2	3	2	2
	2018	1	1	2	2	2	1	2	1
Chongqing	2012	3	2	2	3	2	3	3	3
	2018	3	3	3	3	2	3	3	3
Sichuan	2012	3	2	2	3	2	3	3	3
	2018	4	3	3	3	2	3	3	3
Guizhou	2012	3	2	2	3	2	3	2	2
	2018	1	3	2	2	2	2	1	2
Yunnan	2012	3	2	2	3	2	3	2	2
	2018	2	1	2	2	2	2	1	2
Shaanxi	2012	3	2	2	3	2	3	3	2
	2018	4	3	2	3	2	3	3	2
Gansu	2012	3	2	2	3	2	3	2	2
	2018	1	1	2	1	2	1	1	2



TABLE 4: Continued.

		(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
Qinghai	2012	1	1	1	1	1	1	1	1
	2018	1	1	1	1	1	1	1	1
Ningxia	2012	3	1	2	3	2	3	2	2
	2018	1	1	2	1	2	1	1	2
Xinjiang	2012	3	2	2	3	2	3	2	2
	2018	1	1	2	1	2	1	1	2

Note: the input of important innovation factors is divided into group 1 before the first turning point, into group 2 between the first turning point and the maximum slope point, into group 3 between the maximum slope point and the second turning point, and into group 4 after the second turning point.

TABLE 5: Grouping characteristics of important innovation factor inputs.

Grouping characteristics	2012	2018
(a) Part of important innovation factor inputs has not yet reached the starting point of increasing marginal effect	Qinghai, Ningxia	Shanxi, Inner Mongolia, Heilongjiang, Jiangxi, Guangxi, Hainan, Guizhou, Yunnan, Gansu, Qinghai, Ningxia, Xinjiang
(b) All the important innovation factor inputs are in the marginal effects from increasing to maximum and then to falling flat stage	Tianjin, Hebei, Shanxi, Inner Mongolia, Liaoning, Jilin, Heilongjiang, Anhui, Fujian, Jiangxi, Henan, Hubei, Hunan, Guangxi, Hainan, Chongqing, Sichuan, Guizhou, Yunnan, Shaanxi, Gansu, Xinjiang	Tianjin, Hebei, Liaoning, Jilin, Anhui, Henan, Hunan, Chongqing
(c) Part of important innovation factors has completed the input required for the optimal combination of innovation factors	Beijing, Shanghai, Zhejiang, Shandong	Beijing, Shanghai, Jiangsu, Zhejiang, Fujian, Shandong, Hubei, Sichuan, Shaanxi
(d) All the important innovation factor inputs meet the requirements of optimal combination	Jiangsu, Guangdong	Guangdong

innovation output. There are also some provinces where the investment in important innovation factors has risen from the characteristic (b) in 2012 to the characteristic (c) in 2018, such as Fujian, Hubei, Sichuan, and Shaanxi. The grouping characteristics have been upgraded, and the above provinces need to continue to maintain this situation, or they can increase investment in innovation factors which have a large gap such as trade freedom and other factors. Thirdly, the input of important innovation factors in some provinces in 2012 and 2018 has always remained at the characteristic (c), such as Beijing, Shanghai, Zhejiang, and Shandong. The above provinces may be based on the gap of some important innovation factors between current and optimal inputs, combined with the contribution of important innovation factor inputs, giving priority to the important innovation factors with small gaps and large contributions. Fourthly, the inputs of important innovation factors in Guangdong in 2012 and 2018 have met the requirements of characteristics (d). It means that Guangdong has realized the effective allocation of innovation factors in 2018. If they want to improve the optimal innovation output standard, it is more difficult to start from the input end of innovation factors. It may be achieved by changing the existing innovation development patterns. The important innovation factor inputs in Jiangsu Province have declined from the characteristics (d) in 2012 to the characteristics (c) in 2018. Jiangsu Province may continue to increase inputs in financial development and data applications in the future.

### 5. Conclusions and Future Research

Innovation is the driving force of economic development, and effectively measuring the allocation of innovation factors is conducive to better playing the role of innovation. We use the random forest method to identify the effectiveness of innovation factor allocation in 30 provinces in China from 2009 to 2018 and explore the importance of the innovation factor inputs' contribution in the entire sample period and its own maximum contribution timing. Then we further grasp the marginal effect characteristics of important innovation factor inputs, catch the time changes of its key boundary points, analyze the actual inputs of each province based on the optimal combination of important innovation factor inputs, and finally propose future adjustment directions. The specific conclusions reached are as follows.

First, more than half of the provinces is not fully effective in the innovation factor allocation. How to rely on internal and external innovation factors and make use of institution and data innovation factors to achieve the actual innovation output higher than their own optimal levels is still in a period of exploration and adjustment. Second, from the perspective of the importance of innovation factor inputs, data innovation factors especially data integration and data applications have contributed significantly, which reflects the important impact of data innovation factors on innovation output. From the view of the maximum contribution time of innovation factor input, the most innovation factor inputs

contribute its own maximum value in 2018, suggesting that the structure of China's innovation factor input has an overall upward development trend. Third, the marginal effects of important innovation factor inputs are nonlinear, and the positions of the first turning point, the maximum slope point, and the second turning point of the marginal effect of important innovation factor inputs in different periods are different, indicating that the marginal effect of important innovation factor inputs on innovation output is not always the same. Based on this, we find out the optimal combination of important innovation factor inputs. Fourth, Guangdong Province has achieved the optimal combination of important innovation factor inputs or can change the existing innovation development pattern, while the remaining provinces have not yet reached the optimal factor input requirements and need to adjust the inputs to improve the current innovation choice.

Based on the above conclusions, the following policy implications can be obtained. Firstly is improving the effectiveness of the innovation factor allocation further. We not only need to pay attention to the coordination of internal and external innovation factors, but also focus on the coupling of institution and data innovation factors. Secondly is giving full play to the potential of data innovation factors, especially data application and data integration: Developed provinces may be able to make further development, while backward provinces may be able to achieve curve overtaking. Thirdly is grasping the marginal effects of important innovation factor inputs, and making good use of the impact trend of the important innovation factor inputs. Fourth, each province makes appropriate innovation development policies based on local conditions and rationally adjusts the current innovation factor inputs selection according to the optimal combination.

Needless to say, this paper has some limitations, which are also the basis of future research in this paper. Firstly, when evaluating the effectiveness of innovation factors allocation, this paper takes internal innovation factors, external innovation factors, institutional innovation factors, and data innovation factors as innovation factor inputs, and per capita new product sales revenue as innovation output. Except for per capita new product sales revenue, it can also be measured by indicators such as per capita gross national product in the evaluation of innovation output. Therefore, it is necessary to add other variables that can measure innovation output to identify the effectiveness of innovation factor allocation in further research, so as to more comprehensively show the allocation situation of innovation factors in various provinces in China.

Secondly, this paper takes China's provinces as the research object and evaluates the innovation factor allocation from the macrolevel. In addition, it also has practical significance of evaluating the innovation factor allocation from the meso- and microlevels. Therefore, it is necessary to analyze the innovation factor allocation from the industry level and the enterprise level and further explore the future adjustment direction of the innovation factor allocation at the meso- and microlevels.

Finally, when analyzing the impact of innovation factor input, this paper focuses on the marginal effect trend and key boundary points of important innovation factor input but does not show the impact of remaining innovation factor input on innovation output. Although the impact of this part of innovation factor inputs is not significant, they also make the marginal contributions to innovation output. In the future, we may further explore the marginal effect characteristics and key boundary points of such innovation factor inputs. It may draw other important conclusions.

## Data Availability

This paper is based on the panel data of 30 provinces in China from 2009 to 2018 (except Tibet). The data sources and processing methods are shown in Table 1.

## Conflicts of Interest

The authors declare no conflicts of interest.

## Authors' Contributions

Mo designed the computational frame work and wrote the manuscript. Y. W. designed the model and analyzed the data. C. Q. conducted empirical research. S. F. collected the relevant literature and determined the indicator. All authors discussed the results and contributed to the final manuscript. All authors have read and agreed to the published version of the manuscript.

## Acknowledgments

This research was supported by the characteristic and preponderant discipline of Key Construction Universities in Zhejiang Province (Zhejiang Gongshang University, Statistics), the National Natural Science Foundation of China (71763010 and 71803038), and Jiangxi Provincial Social Science "14th Five-Year" Plan Project (21YJ05).

## References

- [1] S. Duan, T. W. Jiang, J. Y. Zhang, and J. L. Wang, "Research on technological innovation evaluation of regional enterprises: analysis of technological innovation evaluation indicators for Zhejiang's enterprises," *China Soft Science*, vol. 29, no. 5, pp. 85–96, 2014.
- [2] Z. M. Dai, W. T. Zheng, and B. B. Yang, "Research on the regional differences of China's science and technology financial efficiency from the perspective of scope scale," *Studies in Science of Science*, vol. 35, no. 9, pp. 1326–1333, 2017.
- [3] C. Q. Tao and Y. Z. Peng, "Spatial spillover effect of knowledge potential energy on regional technological innovation efficiency under the institutional proximity," *Contemporary Finance & Economics*, vol. 39, no. 2, pp. 15–25, 2018.
- [4] Y. Li, X. H. Dang, X. F. Han, and W. F. Song, "The study on heterogeneity effect of environmental regulation's long-term & short-term influence on technology innovation——two

- stage analysis based on value chain perspective,” *Studies in Science of Science*, vol. 32, no. 6, pp. 937–949, 2014.
- [5] J. Wang and M. Lan, “Reform of mixed ownership and innovation efficiency of state-owned enterprise—an SNA-based analysis,” *Statistical Research*, vol. 36, no. 11, pp. 90–103, 2019.
- [6] B. Chen and T. Li, “Do fiscal decentralization and environmental regulation promote green technology innovation in China?” *Statistical Research*, vol. 37, no. 6, pp. 27–39, 2020.
- [7] D. Dollar and S. Wei, *Das (Wasted) Kapital: Firm Ownership and Investment Efficiency in China*, p. 13103, IMF Working Papers, Washington, D. C., 2007.
- [8] C.-T. Hsieh and P. J. Klenow, “Misallocation and manufacturing TFP in China and India,” *Quarterly Journal of Economics*, vol. 124, no. 4, pp. 1403–1448, 2009.
- [9] Y. W. Chen and W. M. Hu, “Distortions, misallocation and losses: theory and application,” *China Economic Quarterly*, vol. 10, no. 4, pp. 1401–1422, 2011.
- [10] L. Berchicci, “Towards an open R&D system: internal R&D investment, external knowledge acquisition and innovative performance,” *Research Policy*, vol. 42, no. 1, pp. 117–127, 2013.
- [11] J. J. Hong and L. J. Shi, “Independent R&D, region-specific institutions and innovation performance: evidence from 371 innovative enterprises in China,” *Studies in Science of Science*, vol. 35, no. 2, pp. 310–320, 2017.
- [12] H. Chen and J. Hou, “Independent R&D innovation, knowledge accumulation and science-technology performance: the dynamic threshold mechanism based on high-tech industry,” *Studies in Science of Science*, vol. 34, no. 9, pp. 1301–1309+1425, 2016.
- [13] M. Kafourous, C. Wang, P. Piperopoulos, and M. Zhang, “Academic collaborations and firm innovation performance in China: the role of region-specific institutions,” *Research Policy*, vol. 44, no. 3, pp. 803–817, 2015.
- [14] G. Asimakopoulos, A. J. Revilla, and K. Slavova, “External knowledge sourcing and firm innovation efficiency,” *British Journal of Management*, vol. 31, no. 1, pp. 123–140, 2020.
- [15] L. Nieves and D. Petra, “The interaction between external and internal knowledge sources: an open innovation view,” *Journal of Knowledge Management*, vol. 18, no. 2, pp. 430–446, 2014.
- [16] D. J. Hao and R. Zhang, “Factor endowment, agglomeration effect and the transformation of economic growth momentum,” *Economist*, vol. 30, no. 1, pp. 41–49, 2018.
- [17] J. Schumpeter, *Theory of Economic Development*, Harvard University Press, Cambridge, MA, 1934.
- [18] M. Edwards-schachter, “The nature and variety of innovation,” *International Journal of Innovation Studies*, vol. 2, no. 2, pp. 65–79, 2018.
- [19] F. Gault, “Defining and measuring innovation in all sectors of the economy,” *Research Policy*, vol. 47, no. 3, pp. 617–622, 2018.
- [20] M. Dziallas and K. Blind, “Innovation indicators throughout the innovation process: an extensive literature analysis,” *Technovation*, vol. 80-81, pp. 3–29, 2019.
- [21] C. Q. Tao and M. Xu, “Measuring the allocation level of China innovation elements from the perspective of economic high-quality development,” *The Journal of Quantitative & Technical Economics*, vol. 38, no. 3, pp. 3–22, 2021.
- [22] P. F. Zhu and W. M. Xu, “On the impact of government’s S&T incentive policy on the R&D input and its patent output of large and medium-sized industrial enterprises in Shanghai,” *Economic Research Journal*, vol. 49, no. 6, pp. 45–53+94, 2003.
- [23] W. J. Wang and M. X. Gao, “Study on China’s stock of university-industry collaboration,” *The Journal of Quantitative & Technical Economics*, vol. 36, no. 4, pp. 139–154, 2019.
- [24] Z. G. Ouyang and P. Chen, “Factor endowment, local industrial development and selecting industries,” *Economic Research Journal*, vol. 55, no. 1, pp. 82–98, 2020.

## *Retraction*

# **Retracted: Order Management and Completion Date Prediction of Manufacturing Job-Shop Based on Deep Learning**

### **Computational Intelligence and Neuroscience**

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### **References**

- [1] M. Wang, "Order Management and Completion Date Prediction of Manufacturing Job-Shop Based on Deep Learning," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3458116, 10 pages, 2022.

## Research Article

# Order Management and Completion Date Prediction of Manufacturing Job-Shop Based on Deep Learning

**Mei Wang** 

*School of Information Science and Engineering, Shandong Agriculture and Engineering University, Jinan 250100, China*

Correspondence should be addressed to Mei Wang; [z2013462@sdaeu.edu.cn](mailto:z2013462@sdaeu.edu.cn)

Received 24 January 2022; Revised 19 February 2022; Accepted 24 February 2022; Published 16 March 2022

Academic Editor: Daqing Gong

Copyright © 2022 Mei Wang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

To cope with the volatility of customer order demand, enterprises need to formulate a reasonable production plan based on customer demand for the completion period and their current manufacturing capacity. The existing studies have not fully considered the complex processing procedures, the features of manufacturing attributes, and the repetitive orders of stable consumers. To solve these problems, this paper explores the order management and completion date prediction of manufacturing job-shop based on deep learning. Specifically, the features of manufacturing attributes were extracted and used to predict the activities and completion time of different manufacturing tasks in order management. In addition, a deep learning prediction model was constructed based on a bidirectional long short-term memory network (BiLSTM) and self-attention mechanism, which completes the order management and completion date prediction.

## 1. Introduction

With the continuous development of the market economy and the advancement of science and technology, consumers expect ordered products to be delivered within a shorter cycle [1–7]. The increasingly market-centric order production method puts forward extremely strict requirements on the corporate capacity of production and timely completion [8–10]. To cope with the volatility of customer order demand, i.e., meet consumer requirements on the quality and completion date of ordered products, enterprises need to effectively control the ordered tasks in the manufacturing job-shop and to formulate a reasonable production plan based on customer demand for the completion period and their current manufacturing capacity [11–18]. In the manufacturing industry, order management runs through the entire production cycle. The accurate prediction of the product completion period is the main factor affecting the decision of order management and control [19–22].

Chen [23] constructed a system based on the knowledge of fuzzy neural networks, aiming to improve the performance of manufacturing job-shop in predicting completion

time and allocating internal delivery time. In the system, multiple experts construct their own fuzzy multiple linear regression (MLR) models and predict the job completion time/cycle. Drawing on concepts like machine learning, evolution, and metaheuristic learning, Patil [24] developed an enhanced differentiable dynamic quantization (DDQ) model based on an artificial neural network (ANN). Computational experiments show that the model outshines traditional ANN-based DDQ in the prediction of the completion date, in different job-shop environments and on different volumes of training data. To prevent overfitting from weakening the generalization ability of a single neural network, Zhu et al. [25] introduced a neural network ensemble to propose a Bagging approach based on the cluster analysis of the 0.632 prediction error and conducted a case study to illustrate the whole steps to predict the product due date by using neural network ensemble. The operation of manufacturing job-shop is difficult to manage, owing to the heterogeneity of raw materials, complex transformation process, and varied production flows. Dumetz et al. [26] provided a simulation framework enabling the comparison and evaluation of different production planning strategies and order management strategies. The framework integrates

a basic enterprise resource planning (ERP) system. The user can configure the production plan and order management process and evaluate the model performance in various market environments using the discrete event simulation model. After setting up a set of candidate features, Liu et al. [27] presented a feature selection algorithm based on the self-organizing map-feature-weighted fuzzy c-means (SOM-FWFCM) algorithm. Taking the production data of a job-shop as an example, the proposed algorithm was compared with four feature selection algorithms. The comparison demonstrates the effectiveness of that algorithm.

The completion time of ordered products can be predicted well by combining data mining with the analysis of the discrete data in the manufacturing industry, which are featured by scattered distribution, large volume, and poor authenticity. However, the existing studies have not fully considered the complex processing procedures, the features of manufacturing attributes, and the repetitive orders of stable consumers. To solve these problems, this paper explores the order management and completion date prediction of manufacturing job-shop based on deep learning. Section 2 extracts the features of manufacturing attributes, using the recursive feature elimination approach of random forest (RF), principal component analysis (PCA), and k-means clustering (KMC). Based on the extracted features, Section 3 predicts the activities and completion time of different manufacturing tasks in order management and constructs a deep learning prediction model based on bidirectional long short-term memory network (BiLSTM) and self-attention mechanism. The proposed model was proved effective through experiments.

## 2. Feature Extraction

In complex orders, the diverse products are manufactured in small batches through complicated and variable operations. As a result, it is not very desirable to predict the completion period of the products in complex orders by optimizing the processing operations of these products using heuristic algorithms. In the manufacturing job-shop, the order completion time is affected by complex, stochastic, and correlated factors. During real-world job-shop production, there is no ability to collect all the valuable and real information from the manufacturing process of ordered products. Therefore, it is particularly important to clarify the association between the information containing lots of abnormal data and the completion period of ordered products.

The completion date of ordered products is affected by the following factors: requirement on order quality, urgency of delivery, importance of consumers, profit margin of order, task volume of order, and complexity of operations. This paper proposes a hybrid algorithm for mining the factors affecting the completion period of ordered products. Firstly, the important features of manufacturing attributes were extracted by the recursive feature elimination approach of RF. Then, the linear features of manufacturing attributes were extracted through the PCA. After that, the KMC was applied to extract the nonlinear features of manufacturing

attributes. Finally, the extracted important features, linear features, and nonlinear features are fully fused.

**2.1. Extraction of Important Features.** The recursive feature elimination approach of RF is detailed as follows:

*Step 1.* Perform random sampling with the replacement on the manufacturing attribute samples in the original training set  $C$ . Suppose there are  $M$  original samples, and  $N$  samples are selected in each sampling. Denote the  $i$ -th bootstrap sample set generated through multiple random repetitive sampling as  $\delta_i$ .

*Step 2.* Select the splitting feature and the splitting point with the smallest Gini index to split the decision tree (DT), and build the  $i$ -th nonpruned and fully grown classification and regression tree (CART)  $\psi_i$  based on the bootstrap sample set  $\delta_i$ . Repeat this step until all DTs are constructed.

*Step 3.* Calculate the mean square error (MSE) of the RF model. For the  $i$ -th DT  $\psi_i$ , define the  $M-N$  samples not collected in the  $i$ -th random sampling as a set  $NDS_i$ . Let  $E_i^{\text{MSE}}$  be the MSE of DT  $\psi_i$ ,  $w_c$  be the true value of the test data, and  $w_{ic}^*$  be the predicted value of DT  $\psi_i$ . Then, the MSE can be calculated by the following:

$$E_i^{\text{MSE}} = \frac{\sum_{c=1}^{M-N} (w_c - w_{ic}^*)^2}{M - N}. \quad (1)$$

*Step 4.* Calculate the importance score of each feature of manufacturing attributes. The importance score  $\text{FR}_c$  of feature  $\text{FE}_c$  can be calculated by the following:

$$\text{FR}_c = \frac{1}{l} \sum_{i=1}^l (S_i^{\text{MSE}} - S_i^{\text{MSE}}). \quad (2)$$

*Step 5.* Output the feature set of manufacturing attributes corresponding to the minimum  $E_i^{\text{MSE}}$ , during the elimination of the feature with the smallest  $\text{FR}_c$ .

**2.2. Extraction of Linear Features.** Let  $C = \{c_1, c_2, \dots, c_t\}$  be the original training set composed of the column vectors of manufacturing attribute samples. In this paper, the PCA is adopted to extract the linear features from manufacturing attributes. In essence, the extraction searches for a unit projection vector  $\varphi$  that maximizes the projected variance of  $c_i$  on  $\varphi$ .

To eliminate the influence of the varied dimensionality between the column vectors of manufacturing attribute samples,  $C$  can be normalized by the following:

$$Q = \{q_1, q_2, \dots, q_t\} = \{c_1 - e, c_2 - e, \dots, c_t - e\}, \quad (3)$$

where

$$e = \frac{1}{t} \sum_{i=1}^t c_i. \quad (4)$$

Let  $\Lambda$  be the covariance matrix of sample set  $Q$ . The projection variance  $PV(q)$  of dataset  $Q = \{q_1, q_2, \dots, q_t\}$  on unit vector  $\varphi$  can be calculated by the following:



$$\begin{aligned}
FR(q) &= \frac{1}{t} \sum_{i=1}^t (q_i^\psi \phi)^2 = \frac{1}{t} \sum_{i=1}^t (q_i^\psi \phi)^\psi (q_i^\psi \phi) \\
&= \frac{1}{t} \sum_{i=1}^t \phi^\psi q_i q_i^\psi \phi = \phi^\psi \left( \frac{1}{t} \sum_{i=1}^t q_i q_i^\psi \right) \phi.
\end{aligned} \tag{5}$$

Thus, the PCA can be converted into the following mathematical problem:

$$\begin{cases} \max_{\phi} \phi^\psi \sum \phi \\ \text{s.t. } \phi^\psi \phi = 1 \end{cases}. \tag{6}$$

Let  $\mu$  be the Lagrangian multiplier. Formula (6) can be solved by the Lagrangian function  $U(\phi, \mu)$ :

$$U(\phi, \mu) = \phi^\psi \sum \phi - \mu \phi^\psi \phi. \tag{7}$$

Solving the partial derivative of  $U(\phi, \mu)$  for  $\phi$ :

$$\frac{\partial U(\phi, \mu)}{\partial \phi} = 2 \sum \phi - 2\mu \phi = 0 \Rightarrow \sum \phi = \mu \phi. \tag{8}$$

Combining  $\Lambda \phi = \mu \phi$  with formula (8):

$$FR(q) = \phi^\psi \left( \frac{1}{t} \sum_{i=1}^t q_i q_i^\psi \right) \phi = \phi^\psi \sum \phi = \mu \phi^\psi \phi = \mu. \tag{9}$$

Formula (9) shows that the eigenvalue of the covariance matrix  $\Lambda$  of manufacturing attribute sample  $q$  is its projection variance on unit vector  $\phi$ . The largest eigenvalue of  $\Lambda$  is the maximum projection variance, and the second largest eigenvalue of  $\Lambda$  is the next best projection variance. The rest can be obtained by analogy. The flow of the PCA is detailed as follows:

*Step 1.* Normalize the original training set  $C = \{c_1, c_2, \dots, c_t\}$  to eliminate the influence of the varied dimensionality between the column vectors of manufacturing attribute samples. That is, ensure that the mean projection  $e'$  of normalized dataset  $Q = \{q_1, q_2, \dots, q_t\}$  on projection vector  $\phi$  equals zero:

$$e' = \frac{1}{t} \sum_{i=1}^t q_i^\psi \phi = \left( \frac{1}{t} \sum_{i=1}^t q_i^\psi \right) \phi = 0. \tag{10}$$

*Step 2.* Compute the covariance matrix  $\Lambda$  of sample set  $Q$ :

$$\Lambda = \frac{1}{t} \sum_{i=1}^t q_i q_i^\psi. \tag{11}$$

*Step 3.* Decompose the eigenvalues of covariance matrix  $\Lambda$ , and sort them in descending order. Denote the ranked eigenvalues and the corresponding eigenvectors as  $\{\xi_1, \xi_2, \dots, \xi_t\}$  and  $\{\varphi_1, \varphi_2, \dots, \varphi_t\}$ , respectively.

*Step 4.* Map the eigenvectors  $\{\varphi_1, \varphi_2, \dots, \varphi_\nu\}$  corresponding to the top  $\nu$  eigenvalues. Let  $q_i^N$  be the mapped sample data. Then, the  $t$ -dimensional data can be mapped into  $\nu$ -dimensional data by the following:

$$q_i^{\text{new}} = \begin{bmatrix} \phi_1^\psi q_i \\ \phi_2^\psi q_i \\ \vdots \\ \phi_\nu^\psi q_i \end{bmatrix}. \tag{12}$$

The cumulative contribution rate  $\gamma$  can be determined by the following:

$$\gamma = \sqrt{\frac{\sum_{i=1}^{\nu} \xi_i^2}{\sum_{i=1}^t \xi_i^2}}. \tag{13}$$

*2.3. Extraction of Nonlinear Features.* The KMC was adopted to extract the nonlinear features of manufacturing attributes. Before feature extraction, the low-dimensional data on ordered products should be mapped into high-dimensional data through the following steps:

*Step 1.* Perform feature interaction on the original training set  $C = \{c_1, c_2, \dots, c_t\}$  and denote the resulting interactive feature set as  $C' = \{c_{0,1}, c_{1,2}, \dots, c_{t-1,t}\}$ . Let  $q_{ij}$  be the area feature interaction result between features  $q_i$  and  $q_j$ . The feature product interaction can be calculated by the following:

$$q_{i,j} = q_i \times q_j. \tag{14}$$

*Step 2.* Set the number of clusters for nonlinear manufacturing features and select an initial center randomly for each class. Find the optimal number of classes by the inflection point method for the within-cluster sum of squares, i.e., compute the within-cluster sum of squares at different  $k$  values. Let  $SU(f_1, f_2, \dots, f_K)$  be the within-cluster sum-of-squares;  $f_j$  be the center of the  $j$ -th cluster;  $q_i$  be the  $i$ -th sample in the  $j$ -th cluster;  $r_j$  be the total number of samples in the  $j$ -th cluster. Then, the optimal number of clusters can be calculated by the following:

$$SU(f_1, f_2, \dots, f_K) = \sum_{j=1}^K \sum_i^{r_j} (q_i - f_j)^2. \tag{15}$$

*Step 3.* Compute the distance between each cluster center and each interactive feature in the dataset. Based on the computed results, assign the interactive feature to the right cluster. Measure that distance with cosine similarity. The cosine similarity between samples  $q_1$  and  $q_2$  can be calculated by the following:

$$v = \frac{\sum_{i=1}^r q_{1i} q_{2i}}{\sqrt{\sum_{i=1}^r q_{1i}^2} \sqrt{\sum_{i=1}^r q_{2i}^2}} \tag{16}$$

*Step 4.* Update the center of each cluster based on the mean of the samples in that cluster. The partial derivative of the loss function can be calculated by the following:

$$\begin{aligned} \frac{\partial S U(f_1, f_2, \dots, f_h)}{\partial f_j} &= \sum_{j=1}^h \sum_i^{r_j} \frac{\partial (q_i - f_j)^2}{\partial f_j} \\ &= \sum_i^{r_j} \frac{\partial (q_i - f_j)^2}{\partial f_j} = \sum_i^{r_j} -2(q_i - f_j). \end{aligned} \quad (17)$$

Making formula (17) equal to zero:

$$\begin{aligned} \sum_i^{r_j} -2(q_i - f_j) &= 0, \\ r_j f_j - \sum_i^{r_j} q_i &= 0 \Rightarrow f_j = \frac{\sum_{i=1}^{r_j} q_i}{r_j}. \end{aligned} \quad (18)$$

Formula (18) shows that the loss is minimized, when the cluster center equals the mean of all samples in that cluster.

*Step 5.* Repeat Steps 3-4 until the termination condition is satisfied.

*Step 6.* Output the  $K$  cluster centers as the nonlinear features extracted from manufacturing attribute samples.

Figure 1 explains the flow of linear and nonlinear feature extractions. All completion date prediction models are verified on the test data. The above algorithm is adopted to extract the features of the factors affecting the completion date, according to the test error of each model. Then, the mean test error is computed, and the prediction model with the smallest test error is selected.

### 3. Prediction of Activities and Completion Time

Process mining of predictive order management helps job-shop managers identify abnormal, incompliant order management activities so that they could take emergency measures and crisis response measures. In the context of order management, the prediction of activities and completion time of each manufacturing task brings several advantages: effectively increasing job-shop production efficiency, significantly lowering operating costs, and accurately recognizing illegal activities.

Figure 2 presents the framework of the prediction model for order management activities and completion time. There are five layers in the framework: an input layer, an embedding layer, a BiLSTM layer, a self-attention layer, and an output layer.

The attributes in job-shop manufacturing logs should be extracted and converted into eigenvectors before being imported to the prediction model. Let  $K = \{\varepsilon_1, \varepsilon_2, \varepsilon_3, \dots, \varepsilon_m\}$  be the set of job-shop manufacturing logs, where  $\varepsilon_i = \langle o_{\tau 1}, o_{\tau 2}, o_{\tau 3}, \dots, o_{\tau m} \rangle$  represents the evolutionary trajectory of the  $i$ -th event, with  $m = |\varepsilon_i|$ . Each evolution trajectory needs to be converted into an eigenvector  $a = [a_1, a_2, a_3, \dots, a_{GR}]$ , with  $GR$  being the number of samples. In the eigenvector  $a =$

$[a_1, a_2, a_3, \dots, a_{GR}]$ , element  $a_i$  is a two-dimensional eigenvector, including both the event trajectory related to completion time and the number of event-related attributes. The latter covers numerical and non-numerical attributes.

Before an eigenvector is inputted to the neural network via the embedding layer, it is necessary to linearly map the high-dimensional sparse eigenvector  $a = [a_1, a_2, a_3, \dots, a_{GR}]$ . Let  $WS$  be the dimensionality of the code for the mapped eigenvector. Then, the mapped low-dimensional dense eigenvector can be expressed as  $o = [o_1, o_2, o_3, \dots, o_\tau]$ , with  $o_\tau \in R^{WS}$ . Then, we have the following:

$$\begin{aligned} GR(a_1, a_2, a_3, \dots, a_{GR}|a_i) &= \sum_{d=1}^{GR} GR(a_d|a_i), \\ o(a) &= - \sum_{i=1}^k \sum_{i=1}^k \log GR(a_{GR}|a_i). \end{aligned} \quad (19)$$

In our embedding layer, nonnumerical attributes are encrypted by one-hot encoding. This is because one-hot encoding features high-dimensional sparsity and involves no internal association through nonlinear mapping. Every embedding vector can be updated through the network training based on the embedding layer, completing the search for similarity between different vectors in a multi-dimensional space [28].

The BiLSTM consists of two LSTMs with the opposite propagation directions. Figure 3 explains the internal structure of an LSTM. As a modification of recurrent neural network (RNN), the LSTM has a strong ability of modeling time series and overcomes the memory problem and vanishing gradients of traditional RNN. The LSTM can update the hidden state  $YC_\tau$ , based on the input  $o_\tau$  of the previous layer and the previous hidden state  $YC_{\tau-1}$ . Let  $SR_\tau$ ,  $YW_\tau$ , and  $SC_\tau$  be the input gate, forget gate, and output gate, respectively. The input gate selectively preserves the input information and updates the cell state, the forget gate selectively forgets the redundant information, and the output gate determines which cell state should be outputted. Let  $\omega$  be the output weight;  $PO$  be the bias;  $JH_1$  and  $JH_2$  be the activation function sigmoid and the activation function tanh, respectively. Then, we have the following:

$$\begin{aligned} SR_\tau &= JH_1(\omega_{SR} * [YC_{\tau-1}, o_\tau] + PO_{SR}), \\ YW_\tau &= JH_1(\omega_g * [YC_{\tau-1}, o_\tau] + PO_{YW}), \\ GS'_\tau &= JH_2(\omega_{GS} * [YC_{\tau-1}, o_\tau] + PO_{GS}), \\ GS_\tau &= YW_\tau * GS_{\tau-1} + SR_\tau * GS'_\tau, \\ SC_\tau &= JH_1(\omega_{SC} * [YC_{\tau-1}, o_\tau] + PO_{SC}), \\ YC_\tau &= SC_\tau * JH_2(GS_\tau). \end{aligned} \quad (20)$$

The BiLSTM is an extension of unidirectional LSTM based on reverse time flow. Let  $YC'_\tau$  and  $YC''_\tau$  be the forward and backward order management states, respectively. Then, the current output state  $YC_\tau$  can be updated based on  $YC'_\tau$  and  $YC''_\tau$ . In addition, the one-way operation flow of the LSTM is denoted as  $K_{OW}$ . Compared with traditional one-way LSTM, the BiLSTM learns historical and

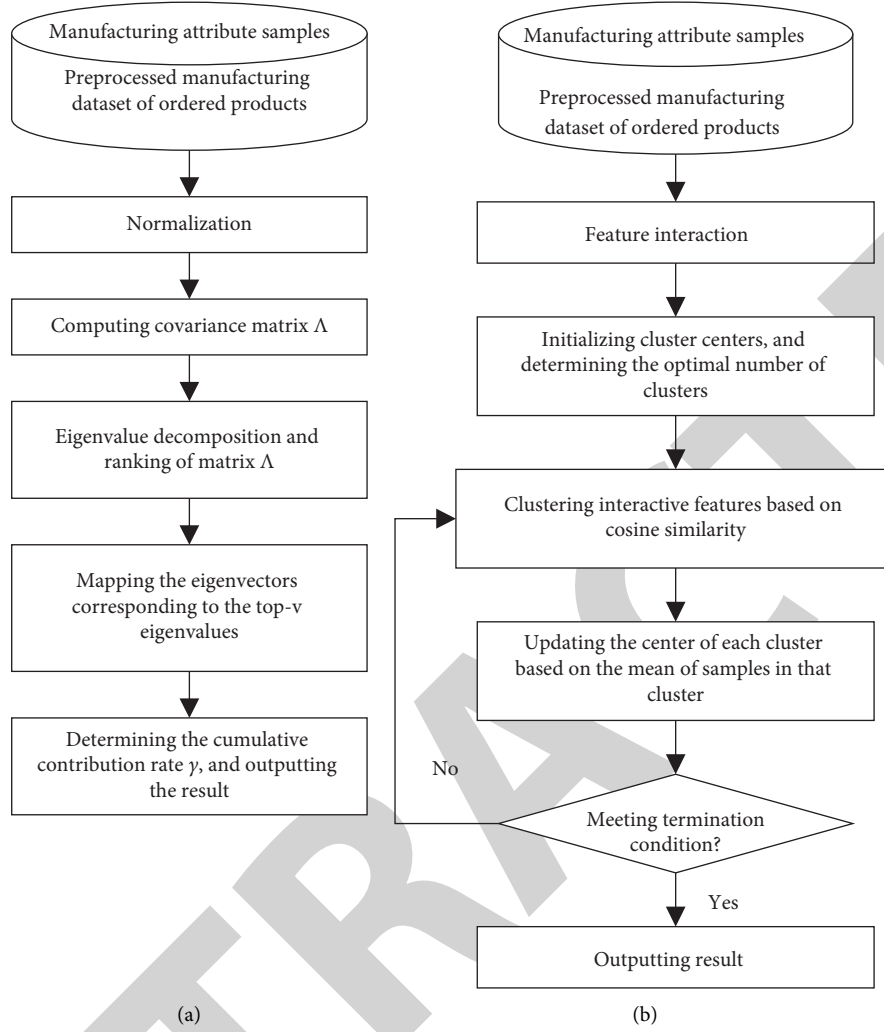


FIGURE 1: Flow of linear and nonlinear feature extractions. (a) Linear feature extraction; (b) nonlinear feature extraction.

future states of order management simultaneously, and acquires highly stable and reliable feature information. Then, we have the following:

$$\begin{aligned}
 YC'_{\tau} &= K_{OW}'(YC_{\tau-1}, o_{\tau}), \\
 YC''_{\tau} &= K_{OW}''(YC_{\tau+1}, o_{\tau}), \\
 YC_{\tau} &= [YC'_{\tau}, YC''_{\tau}].
 \end{aligned} \tag{21}$$

The job-shop orders placed by a stable consumer group tend to be repetitive. In fact, repetition is an obvious feature of predicting the activities and completion time of every manufacturing task during order management. This paper adds a self-attention mechanism to the network to forecast repetitive activities. Figure 4 explains the internal structure of the attention mechanism. The attention-based prediction model considers the weight coefficient between input eigenvectors, and the manufacturing tasks related to the current input eigenvector, without being distracted by the information weakly correlated with the current input eigenvector.

Let  $x_{\tau, \tau'}$  be the similarity matrix between hidden states  $YC_{\tau}$  and  $YC'_{\tau'}$  at moments  $\tau$  and  $\tau'$ , respectively;  $\omega_{YW}$  and

$\omega_{YW}'$  be the weight matrices of hidden states  $YC_{\tau}$  and  $YC'_{\tau'}$ , respectively;  $\omega_x$  be the weight matrix for nonlinear combination;  $PO_x$  and  $PO_{YW}$  be bias vectors. The vector matrix  $[YC_1, YC_2, YC_3, \dots, YC_{\tau}]$  of the previous BiLSTM is imported to the attention layer. Then, the similarity of any feature to every neighbor can be characterized by a self-attention matrix  $X$ :

$$\begin{aligned}
 YW_{\tau, \tau'} &= JH_2(\omega_{YW} YC_{\tau} + \omega_{YW}' YC'_{\tau'} + PO_{YW}), \\
 x_{\tau, \tau'} &= JH(\omega_x YW_{\tau, \tau'} + PO_x).
 \end{aligned} \tag{22}$$

The hidden state  $\tau'$  of the attention at moment  $\tau$  is the weighted sum of all hidden states  $YC'_{\tau'}$  and similarity matrix  $x_{\tau, \tau'}$  at moment  $\tau'$ :

$$k_{\tau} = \sum_{\tau'=1}^m x_{\tau, \tau'} \times YC'_{\tau'}. \tag{23}$$

For each manufacturing task in order management, the output layer should predict both activities and completion time. The activity prediction mainly transforms order management into multiple classes of manufacturing tasks.

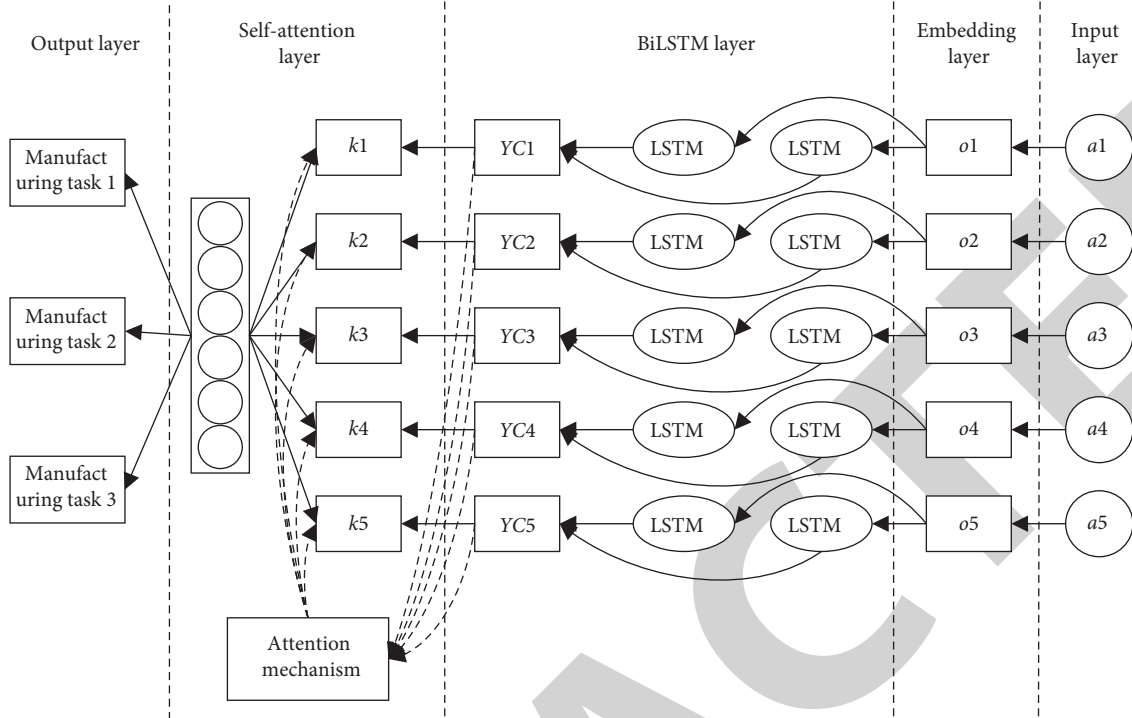


FIGURE 2: Prediction model for order management activities and completion time.

The softmax classifier is adapted to output the predicted activities of order management and the cross-entropy loss. Let  $a$  be the evolution trajectory of the input manufacturing event; FS be the total classes of manufacturing tasks;  $b_i$  be the true label class of the  $i$ -th class;  $g_i(a)$  be the predicted output of the model. Then, the cross-entropy loss can be calculated by the following:

$$\text{LOSS}(a) = - \sum_{i=k}^{FS} b_i \log g_i(a). \quad (24)$$

#### 4. Experiments and Results Analysis

Taking a furniture enterprise in Foshan, southern China's Guangdong Province as an example, this paper adopts the recursive feature elimination approach of RF to extract the important features from manufacturing attributes. The lowest-ranking feature was removed, and the feature extraction error was computed. These steps were repeated until all unimportant features were eliminated. Figure 5 reports the feature extraction errors at different numbers of residual features. It can be learned that the model was highly accurate at 5–7 residual features and relatively inaccurate at fewer than 4 residual features.

In this paper, the PCA is performed to extract the linear features of manufacturing attributes. The cumulative contribution rate was calculated under a different number of principal components (as shown in Figure 6). Referring to the cumulative contribution rates of all components, the minimum threshold was defined for the cumulative contribution rate (0.8). Hence, the top six principal components

were extracted. Table 1 shows the linear features extracted from the manufacturing attributes in view of the factors affecting the order completion period.

Next, the KMC was called to extract the nonlinear features of manufacturing attributes. The optimal number of clusters was determined as 15 by the inflection point and the within-cluster sum-of-square. After that, a cluster analysis was carried out on interactive features. Table 2 shows the nonlinear features extracted by KMC from manufacturing attributes.

The prediction performance of our model on manufacturing activities and completion time for ordered products was evaluated through experiments on a self-designed manufacturing attribute sample set. The sample set was divided into a training set and a test set by the ratio of 3 : 1. One-eighth of the training set was organized as a verification set. Figure 7 shows the training loss curve of our prediction model for the completion period. The curve demonstrates the convergence ability of our model.

Our prediction model was applied to predict the manufacturing activities and completion time on samples from different sources (as shown in Table 3). The prediction accuracy and error of our model were compared with those of stacked autoencoder and 1DCNN. From the manufacturing sample sets from different job-shops, the trajectories of manufacturing events with repeatable features involving orders from stable consumers were screened. In addition, our model was compared separately with the stacked autoencoder. Figure 8 compares their prediction accuracies of repetitive activities on samples from different sources. Facing samples from different sources, our model was far more accurate in predicting repetitive activities than

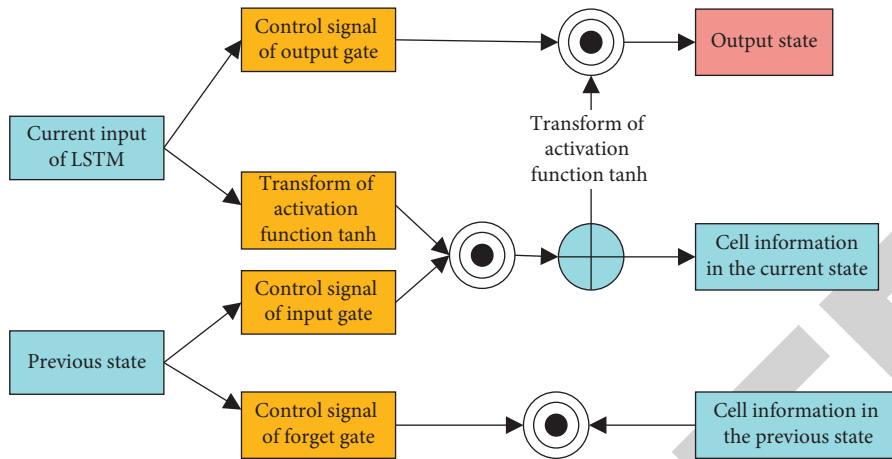


FIGURE 3: Internal structure of an LSTM.

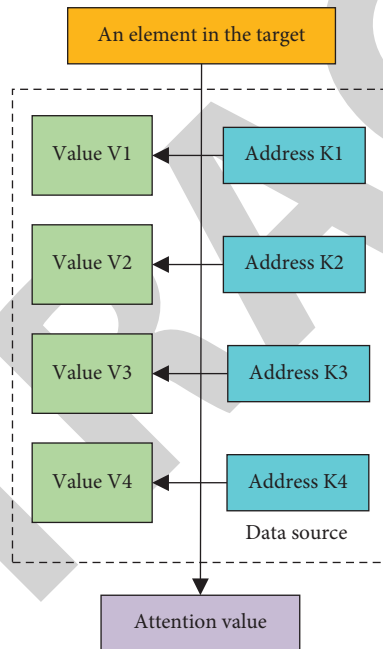


FIGURE 4: Internal structure of attention mechanism.

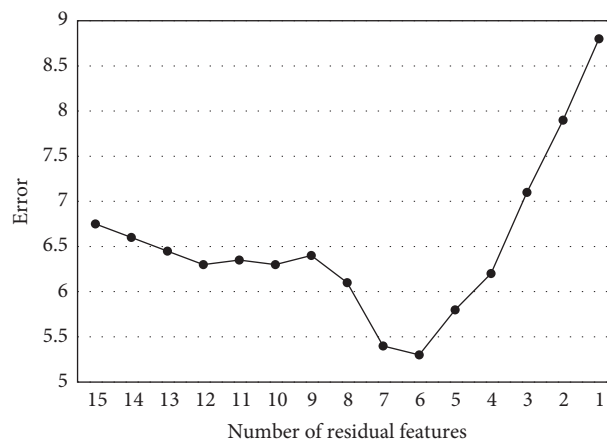


FIGURE 5: Feature extraction errors vs. number of residual features.

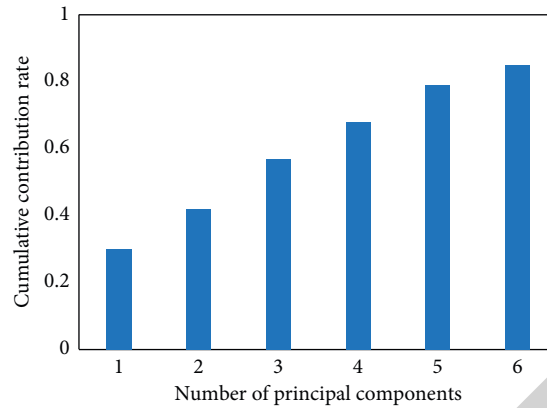


FIGURE 6: Cumulative contribution rate vs. number of principal components.

TABLE 1: Results of linear feature extraction.

Serial number	1	2	3	4	5	6	7	...	20
First principal component	0.8605	-0.9658	0.6122	0.9354	0.3251	0.7523	-0.8695	...	1.1352
Second principal component	0.7258	1.5384	2.0314	0.3528	-1.6583	-2.2587	-2.8391	...	1.0038
Third principal component	0.0358	-0.3623	0.5682	-0.3529	-0.0595	-1.1536	0.1853	...	2.3844
Fourth principal component	2.1059	-1.0568	0.2685	1.7682	1.1582	1.6569	0.3587	...	0.2859
Fifth principal component	-0.2341	-0.9659	0.5623	1.7862	1.1456	1.6852	0.3865	...	0.2752
Sixth principal component	-0.2694	-0.2381	0.6852	1.0586	0.6387	-0.5283	-0.9381	...	-0.8375

TABLE 2: Results of nonlinear feature extraction.

Serial number	1	2	3	4	5	6	7	8	...	20
Feature-1	0.3652	0.3424	0.3976	0.0531	0.0283	0.1385	0.0869	0.1859	...	0.2541
Feature-2	0.6831	0.3157	0.3867	0.2735	0.0586	0.0853	0.1867	0.1527	...	0.3856
Feature-3	0.5672	0.0725	0.5282	0.4685	0.3652	0.8594	0.0685	0.0267	...	0.8659
Feature-4	0.8693	0.9658	0.3254	0.0823	0.0428	0.0725	0.2843	0.3048	...	0.5428
Feature-5	0.2864	0.6859	0.4857	0.3257	0.2527	0.2576	0.0865	0.2375	...	0.1586
Feature-6	0.4586	0.0527	0.2586	0.5861	0.3458	0.3528	0.3157	0.2542	...	0.3425

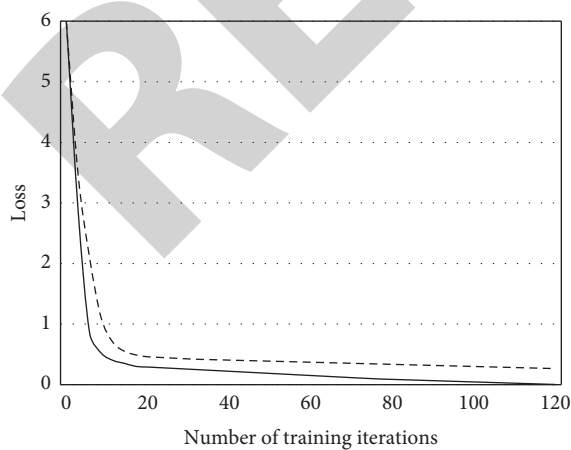
— Our model  
--- Reference model

FIGURE 7: Training loss curve of completion period prediction model.

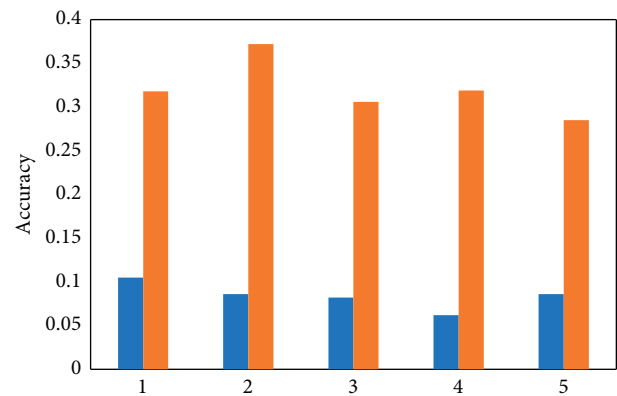
■ Stacked autoencoder  
■ Our model

FIGURE 8: Prediction accuracies of repetitive activities on data from different sources.



TABLE 3: Experimental results of different models.

Source	Job-shop 1			Job-shop 2		
Model type	Our model	Stacked autoencoder	1DCNN	Our model	Stacked autoencoder	1DCNN
Prediction accuracy of manufacturing activities	79.15	74.52	71.51	77.35	76.26	71.38
Prediction error of completion time	1.65	2.41	—	1.56	1.61	—

Note: 1DCNN is short for the one-dimensional convolutional neural network.

TABLE 4: Experimental results of the two prediction models.

Model number		1	2	3	4	5	6	7	8
Feature extraction method		Machine learning	Machine learning	Machine learning	Our model	Our model	Our model	Machine learning	Our model
Test samples	30.25	32.86	29.65	30.159	30.59	30.58	30.84	31.86	30.86
	39.68	34.68	35.67	38.67	39.75	39.64	39.05	40.56	39.57
	35.86	36.28	34.58	36.24	35.67	35.28	35.58	34.28	35.54
	40.58	41.53	39.67	41.25	40.58	40.58	40.91	42.38	40.57
	38.53	37.86	36.82	37.58	38.48	38.82	38.81	37.12	38.58

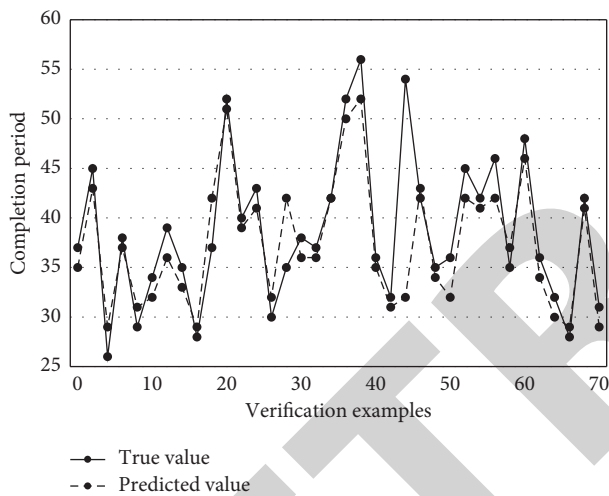


FIGURE 9: Predicted value vs. true value.

the stacked autoencoder. Of course, our model performance is slightly insufficient on the self-designed dataset: the learning ability was not fully exerted, and the obvious features of manufacturing attributes were not ideally extracted. The reason is that the manufacturing events in the job-shops are executed manually, which brings certain stochastic and changeable factors. In spite of that, the experimental results demonstrate that our model is feasible to predict the completion period of orders with repeatable features laid by stable consumers.

Next, our prediction model for the completion period of ordered products, which was constructed based on the features extracted from manufacturing attributes, was compared with machine learning through experiments. Table 4 shows the experimental results of the two prediction models on the test data. The results fully demonstrate the effectiveness of our model. Figure 9 compares the predicted value of our model with the true value. The small error between the two values visually demonstrates the superiority of our model in prediction.

## 5. Conclusions

Based on deep learning, this paper explores the order management and completion date prediction of manufacturing job-shop. Firstly, the important features, linear features, and nonlinear features were extracted from manufacturing attributes. Next, a deep learning prediction model was constructed based on BiLSTM and the self-attention mechanism. Based on the extracted features, the manufacturing task activities and their completion time were forecasted by the proposed model. Through experiments, the relevant features were extracted from manufacturing attributes by the recursive feature elimination approach of RF, PCA, and KMC. The training loss curve of the completion period prediction model was plotted, revealing the convergence ability of the model. In addition, our prediction model was applied to predict manufacturing activities and completion time on samples from different sources. The prediction error and accuracy were summarized. Furthermore, the experimental results of our model were compared with those of stacked autoencoder and 1DCNN. The comparison visually demonstrates the superior prediction effect of our model.

## Data Availability

The data used to support the findings of this study are available from the author upon request.

## Conflicts of Interest

The author declares no conflicts of interest.

## References

- [1] T. Ahmadi, Z. Atan, T. de Kok, and I. Adan, "Optimal control policies for an inventory system with commitment lead time," *Naval Research Logistics*, vol. 66, no. 3, pp. 193–212, 2019.
- [2] T.-H. Nguyen and M. Wright, "Capacity and lead-time management when demand for service is seasonal and lead-

## Research Article

# Research on the Behavior Evolution of Interest Groups in Book Reverse Logistics

Shuihai Dou <sup>1</sup>, Xianyang Su <sup>1</sup>, Yanping Du <sup>1</sup>, Yanming Tang <sup>2</sup>, and Lizhi Peng <sup>1</sup>

<sup>1</sup>Beijing Institute of Graphic Communication, Beijing, China

<sup>2</sup>Zhenkunxing Industry Supermarket (Shanghai) Co., Ltd, Wuhan, China

Correspondence should be addressed to Yanping Du; [duyanping@bigc.edu.cn](mailto:duyanping@bigc.edu.cn)

Received 3 December 2021; Accepted 10 December 2021; Published 16 March 2022

Academic Editor: Daqing Gong

Copyright © 2022 Shuihai Dou et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the gradual expansion of the book logistics market and the year-on-year increase in book publications, the incidence of book reverse logistics continues to increase, and the problem of book companies' inventory backlog has become increasingly prominent. To effectively alleviate the current backlog of book returns and exchanges, this paper constructs a two-party game model of "book publisher-book retailer," analyzes the evolution process of book publishers and book retailers' participation strategies and the influence of parameter changes on stable strategies through theoretical analysis and numerical simulation, and draws the following conclusions. (1) Whether book publishers and book retailers choose to participate in the reverse logistics optimization of book returns and exchanges is closely related to their benefits and costs, and it also depends on whether the other party participates in the reverse logistics optimization of books. (2) When the cost of participating in book reverse logistics reaches a certain condition, the probability of both parties participating in the optimization is the greatest.

## 1. Introduction

With the rapid development of our country's society and economy, people's demand for spiritual culture is increasing day by day. The national inventory of books has increased substantially, reaching 6.64 billion books in 2014 [1]. At the same time, the incidence of reverse logistics in the book industry is the highest, generally around 30% [2]. In a certain period of time, the return rate of best-selling books will exceed 50% or even 60% [2]. Therefore, the problem of book backlog returns is no longer a disturbing factor affecting the economic development of the book industry but has become one of the leading factors that determine the economic development of the book industry. How to establish a good book reverse logistics between book publishers and book retailers to ensure the healthy and orderly development of the book logistics market is one of the key issues faced by managers and scholars in the industry.

To effectively alleviate the reverse logistics problem of book backlog inventory, Ma [3] established a book centralized inventory optimization model on the basis of analyzing the

problems of book reverse logistics. Ning [4] built a collaboration mechanism with the focus on the green supply chain. Morgan [5] linked the sustainable supply chain management, resource input, and operational performance of reverse logistics and analyzed the sustainability of reverse logistics. Faustino et al. [6] introduced the concept of sustainability into reverse logistics and established an applicable reverse logistics process reference model. To accurately analyze the relationship of interest groups when book publishers and book retailers participate or not participate in the optimization of book reverse logistics, this paper consults relevant documents and sorts them out. Yi [7] and Xia [8] established a Stackelberg game model to analyze the group interest behavior of manufacturers and recyclers in reverse logistics. Ye [9] used evolutionary game theory to analyze the relationship between residents, government, and residents in the reverse logistics of waste sorting and recycling. Chen [10] established a game model to analyze the impact of recyclers on the supply chain under different parameter settings of the recycling mode. Wang [11] used game theory to establish a recycling decision model and analyzed product recycling strategies between

manufacturers and retailers. Chen et al. [12], Shi et al. [13], and Ding et al. [14] used game theory to analyze the reverse logistics pricing of goods between manufacturers and recyclers.

Throughout the existing results, scholars at home and abroad have conducted little research on the evolution process and evolution law of the behavior of book reverse logistics interest groups, and they have rarely studied from the perspective of adaptation and evolution in reverse logistics scenarios to study the gradual change and evolution process of the book interest groups' participation strategies and behaviors by factors such as costs and benefits of reverse logistics. This paper takes book publishers and book retailers as research objects, explores the synergy between the two in the book reverse logistics optimization (hereinafter referred to as optimization), analyzes the behavioral evolutionary game of book publishers and book retailers participating in optimization decision making, and conducts numerical simulations to explore ways to effectively improve the active participation of both companies in book reverse logistics optimization, thereby reducing the inventory space occupied by the accumulation of books, reducing the resource waste of books, and making up for part of the vacancy in the study of the behavior evolution of the book reverse logistics interest groups.

## 2. Group Behavior Evolution Model of Book Publishers and Book Retailers

*2.1. Scene Reappearance and Model Assumptions.* In the book reverse logistics, because book publishers and book retailers consider complex and diverse factors when deciding whether to participate in the optimization of book reverse logistics, this paper makes the following assumptions without changing the nature of the problem:

- (1) Participants in the game: book publishers and book retailers are bounded rational subjects, and they will continue to learn and imitate the strategy choices of other individuals and improve their own revenue decisions, so strategies with higher returns will replace strategies with lower returns [15].
- (2) The strategic behavior of the game subject: both book publishers and book retailers can choose a positive strategy or negative strategy. The book publishers' strategy collection is (participation, non-participation), and the book retailers' strategy collection is (participation, non-participation).
- (3) The proportion of behavior strategy: assuming that in the initial stage of the evolutionary game, the proportion of book publishers participating in the optimization is  $X$ , then the proportion of not participating in the optimization is  $1 - X$ ,  $0 \leq X \leq 1$ ; the proportion of book retailers participating in the optimization is  $Y$ ; then, the proportion of not participating in the optimization is  $1 - Y$ ,  $0 \leq Y \leq 1$ .
- (4) Assume that the cost of book publishers choosing to participate in book reverse logistics optimization is

greater than the cost of not participating in book reverse logistics optimization, that is,  $C_1 > S_1$ .

- (5) Assume that the cost of book retailers choosing to participate in book reverse logistics optimization is greater than the cost of not participating in book reverse logistics optimization, that is,  $C_2 > F_2$ .

Based on the above assumptions, this paper constructs a mixed strategy payment matrix between book publishers and book retailers, as shown in Table 1.

The parameters and explanations involved in the hybrid strategy payout matrix are as follows.

- (1)  $\alpha_1$  and  $\alpha_2$ , respectively, represent the production and sales risk ratio of book publishers and book retailers. Production and sales risk refers to all the costs that book publishers and book retailers need to bear due to risks in the process of book reverse logistics, and  $\alpha_1 + \alpha_2 = 1$ .
- (2)  $N$  is all loss costs incurred in the process of book reverse logistics, including invisible losses such as negative impacts on the enterprise. Visible loss costs such as the backlog of unsalable books and excess shipping costs caused by reverse logistics and so on.
- (3)  $C_1$  and  $C_2$ , respectively, represent the cost of book publishers and book retailers choosing to participate in book reverse logistics optimization, including the cost of purchasing new equipment and new technology, and so on.
- (4)  $S_1$  and  $F_2$ , respectively, represent the cost of book publishers and book retailers choosing not to participate in book reverse logistics optimization, including the loading, unloading, handling, sorting, and other costs incurred in the process of book recycling.
- (5)  $B_1$  and  $L_2$ , respectively, represent the revenue of book publishers and book retailers choosing to participate in book reverse logistics optimization.
- (6)  $D_1$  and  $D_2$ , respectively, represent the revenue of book publishers and book retailers choosing not to participate in book reverse logistics optimization.

*2.2. Establishment of Replication Dynamic Equation.* According to the assumptions and the game payment matrix, this paper can get the expected returns  $E_{A1}$  and  $E_{A2}$  of book publishers choosing to participate in the strategy, non-participation strategy, and the average expected return  $E_A$  of book publishers.

$$\begin{aligned}
 E_{A1} &= Y(B_1 - C_1 - \alpha_1 N) + (1 - Y)(D_1 - C_1 - N) \\
 &= Y(B_1 - D_1 + N - \alpha_1 N) + D_1 - C_1 - N, \\
 E_{A2} &= Y(D_1 - S_1 - N) + (1 - Y)(D_1 - S_1 - N) \\
 &= D_1 - S_1 - N, \\
 \bar{E}_A &= XE_{A1} + (1 - X)E_{A2}.
 \end{aligned} \tag{1}$$

Then, we can get its dynamic replication equation  $F(X)$ .

TABLE 1: Matrix of payment.

Game subject		Book retailers	
		Participation (Y)	Non-participation (1 - Y)
Book publishers	Participation (X)	$B_1 - C_1 - \alpha_1 N,$ $L_2 - C_2 - \alpha_2 N$	$D_1 - C_1 - N,$ $D_2 - F_2$
	Non-participation (1 - X)	$D_1 - S_1 - N,$ $D_2 - C_2$	$D_1 - S_1 - N,$ $D_2 - F_2$

$$\begin{aligned}
F(X) &= \frac{dX}{dt} \\
&= X(E_{A1} - \bar{E}_A) \\
&= X(1 - X)(E_{A1} - E_{A2}) \\
&= X(1 - X)[Y(B_1 - D_1 + N - \alpha_1 N) + S_1 - C_1].
\end{aligned} \tag{2}$$

In the same way, this paper can get the expected returns  $E_{B1}$  and  $E_{B2}$  of bookretailers choosing to participate in the strategy, non-participation strategy. The average expected

return  $\bar{E}_B$  of bookretailers, and the replication dynamic equation  $F(Y)$ .

$$E_{B1} = X(L_2 - C_2 - \alpha_2 N) + (1 - X)(D_2 - C_2)$$

$$= X(L_2 - D_2 - \alpha_2 N) + D_2 - C_2,$$

$$E_{B2} = X(D_2 - F_2) + (1 - X)(D_2 - F_2)$$

$$= D_2 - F_2,$$

$$\bar{E}_B = YE_{B1} + (1 - Y)E_{B2}, \tag{3}$$

$$F(Y) = \frac{dY}{dt}$$

$$= Y(E_{B1} - \bar{E}_B)$$

$$= Y(1 - Y)(E_{B1} - E_{B2}),$$

$$= Y(1 - Y)[X(L_2 - D_2 - \alpha_2 N) + F_2 - C_2].$$

Thus, the replication dynamic equation, Jacobian matrix  $J$ , and the corresponding determinants and traces of the game system are obtained. For the sake of discussion below,

let  $\beta_1 = B_1 - D_1 + N - \alpha_1 N + S_1 - C_1$  and  $\beta_2 = L_2 - D_2 - \alpha_2 N + F_2 - C_2$ .

$$\begin{cases} F(X) = \frac{dX}{dt} = X(1-X)[Y(B_1 - D_1 + N - \alpha_1 N) + S_1 - C_1], \\ F(Y) = \frac{dY}{dt} = Y(1-Y)[X(L_2 - D_2 - \alpha_2 N) + F_2 - C_2], \end{cases} \quad (4)$$

$$J = \begin{bmatrix} \frac{\partial F(X)}{\partial X} & \frac{\partial F(X)}{\partial Y} \\ \frac{\partial F(Y)}{\partial X} & \frac{\partial F(Y)}{\partial Y} \end{bmatrix} \quad (5)$$

$$= \begin{bmatrix} (1-2X)[Y(B_1 - D_1 + N - \alpha_1 N) + S_1 - C_1] & X(1-X)(B_1 - D_1 + N - \alpha_1 N) \\ Y(1-Y)(L_2 - D_2 - \alpha_2 N) & (1-2Y)[X(L_2 - D_2 - \alpha_2 N) + F_2 - C_2] \end{bmatrix},$$

$$\det(J) = \begin{vmatrix} \frac{\partial F(x)}{\partial x} & \frac{\partial F(x)}{\partial y} \\ \frac{\partial F(y)}{\partial x} & \frac{\partial F(y)}{\partial y} \end{vmatrix} = \frac{\partial F(x)}{\partial x} \cdot \frac{\partial F(y)}{\partial y} - \frac{\partial F(y)}{\partial x} \cdot \frac{\partial F(x)}{\partial y}, \quad (6)$$

$$= (1-2X)[Y(B_1 - D_1 + N - \alpha_1 N) + S_1 - C_1] \cdot (1-2Y)[X(L_2 - D_2 - \alpha_2 N) + F_2 - C_2] - Y(1-Y)(L_2 - D_2 - \alpha_2 N) \cdot X(1-X)(B_1 - D_1 + N - \alpha_1 N),$$

$$\text{Tr}(J) = \frac{\partial F(x)}{\partial x} + \frac{\partial F(y)}{\partial y} \quad (7)$$

$$= (1-2X)[Y(B_1 - D_1 + N - \alpha_1 N) + S_1 - C_1] + (1-2Y)[X(L_2 - D_2 - \alpha_2 N) + F_2 - C_2].$$

### 3. Analysis on the Stability of Interest Groups Based on Evolutionary Game

This paper analyzes the evolution process of the strategic system of book publishers and book retailers and combines the replication dynamic equation (4) to obtain five evolutionary equilibrium points of system evolution  $(0, 0)$ ,  $(1, 0)$ ,  $(0, 1)$ ,  $(1, 1)$ ,  $(X_0, Y_0)$ , where  $X_0 = (C_2 - F_2)/(L_2 - D_2 - \alpha_2 N)$ ,  $Y_0 = (C_1 - S_1)/(B_1 - D_1 + N - \alpha_1 N)$ . If and only if  $0 < (C_2 - F_2)/(L_2 - D_2 - \alpha_2 N) < 1$ ,  $0 < (C_1 - S_1)/(B_1 - D_1 + N - \alpha_1 N) < 1$ ,  $(X_0, Y_0)$  may be the stable equilibrium point of the system. Since the parameter value range will affect the evolutionary stability strategy of the system, this paper will discuss the stability strategy of the evolutionary system in four cases.

- (1) Case 1:  $\beta_1 > 0, \beta_2 > 0$ . Substitute  $\beta_1 > 0$  and  $\beta_2 > 0$  into formulas (6) and (7) to obtain the expression of the determinant and trace of the system equilibrium point, as shown in Table 2.

To accurately analyze the relationship between book publishers and book retailers participating in the optimization, the parameters are assigned, as shown in Table 3. This paper can get the dynamic evolution path and trend of replication of book

publishers and book retailers in Case 1, as shown in Figure 1.

It can be seen from Case 1 that when  $\beta_1 > 0, \beta_2 > 0$ , that is,  $B_1 - C_1 + N - \alpha_1 N > D_1 - S_1$ ,  $L_2 - C_2 - \alpha_2 N > D_2 - F_2$ , the net revenue of book publishers and book retailers participating in the optimization is greater than the net revenue of not participating in optimization. The two stability policies that exist in the system are related to the initial states of  $X$  and  $Y$ . When both parties participate in a low proportion, the system tends to approach the  $(0, 0)$  strategy, that is, both parties choose not to participate in optimization. When both parties participate in a high proportion of optimization, the system tends to approach the strategy  $(1, 1)$ , that is, both parties choose to participate in optimization.

- (2) Case 2:  $\beta_1 < 0, \beta_2 > 0$ . In the same way, this paper obtains the expression of the determinant and trace of the equilibrium point of the system, as shown in Table 4.

Similarly, the parameters are assigned, as shown in Table 5. This paper can get the dynamic evolution path and trend of replication of book publishers and book retailers in Case 2, as shown in Figure 2.

TABLE 2: The local stability of the equilibrium point in Case 1.

Equilibrium point	det ( $J$ )	tr ( $J$ )	Local stability	Strategy profile
(0, 0)	+	-	ESS	(Non-participation, non-participation)
(1, 0)	+	Uncertain	Unstable point	(Participation, non-participation)
(0, 1)	+	Uncertain	Unstable point	(Non-participation, participation)
(1, 1)	+	-	ESS	(Participation, participation)
$(X_0, Y_0)$	-	0	Saddle point	

TABLE 3: Assignment of each parameter in Case 1.

Parameter	$\alpha_1$	$\alpha_2$	$N$	$C_1$	$C_2$	$S_1$	$F_2$	$B_1$	$L_2$	$D_1$	$D_2$
Assignment	0.5	0.5	20	17	24	10	11	15	50	10	20

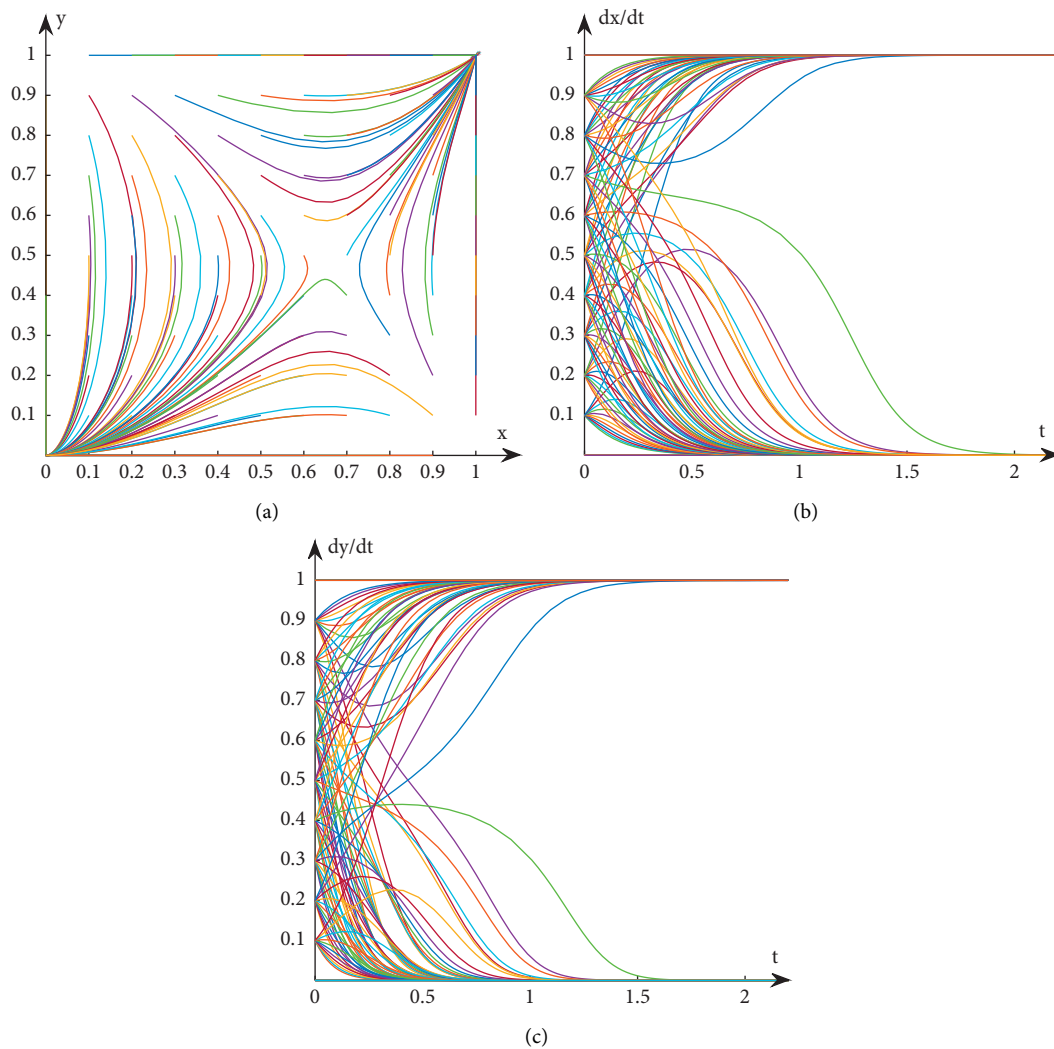


FIGURE 1: Evolution path diagram in Case 1. (a) Evolution path of system. (b) Evolution path of book publishers. (c) Evolution path of book retailers.

(3) Case 3:  $\beta_1 > 0, \beta_2 < 0$ . In the same way, this paper obtains the expression of the determinant and trace of the equilibrium point of the system, as shown in Table 6.

Similarly, the parameters are assigned, as shown in Table 7. This paper can get the dynamic evolution path and trend of replication of book publishers and book retailers in Case 3, as shown in Figure 3.



TABLE 4: The local stability of the equilibrium point in Case 2.

Equilibrium point	$\det(J)$	$\text{tr}(J)$	Local stability	Strategy profile
(0, 0)	+	-	ESS	(Non-participation, non-participation)
(0, 1)	-	Uncertain	Saddle point	(Participation, non-participation)
(1, 0)	+	+	Unstable point	(Non-participation, participation)
(1, 1)	-	Uncertain	Saddle point	(Participation, participation)
$(X_0, Y_0)$	+	0	Unstable point	

TABLE 5: Assignment of each parameter in Case 2.

Parameter	$\alpha_1$	$\alpha_2$	$N$	$C_1$	$C_2$	$S_1$	$F_2$	$B_1$	$L_2$	$D_1$	$D_2$
Assignment	0.5	0.5	20	22	24	4	11	15	45	10	20

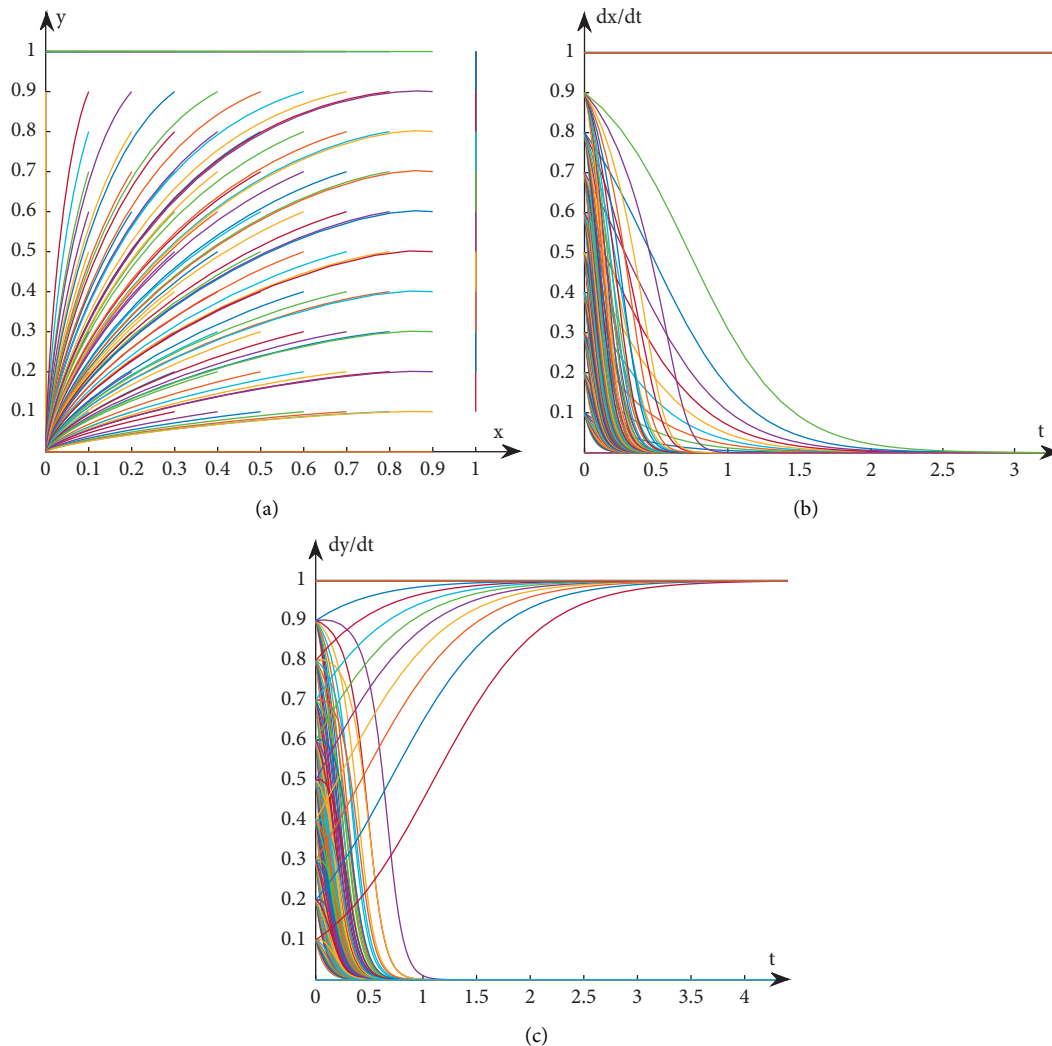


FIGURE 2: Evolution path diagram in Case 2. (a) Evolution path of system. (b) Evolution path of book publishers. (c) Evolution path of book retailers.

It can be seen from Case 2 and Case 3 that when  $\beta_1 < 0$ ,  $\beta_2 > 0$  or  $\beta_1 > 0$ ,  $\beta_2 < 0$ , that is,  $B_1 - C_1 + N - \alpha_1 N < D_1 - S_1$ ,  $L_2 - C_2 - \alpha_2 N > D_2 - F_2$  or  $B_1 - C_1 + N - \alpha_1 N > D_1 - S_1$ ,  $L_2 -$

$C_2 - \alpha_2 N < D_2 - F_2$ , the net revenue of either book publishers or book retailers participating in optimization is less than the net revenue of not participating in the optimization, both parties will not

TABLE 6: The local stability of the equilibrium point in Case 3.

Equilibrium point	$\det (J)$	$\text{tr} (J)$	Local stability	Strategy profile
(0, 0)	+	-	ESS	(Non-participation, non-participation)
(0, 1)	+	+	Unstable point	(Participation, non-participation)
(1, 0)	-	Uncertain	Saddle point	(Non-participation, participation)
(1, 1)	-	Uncertain	Saddle point	(Participation, participation)
$(X_0, Y_0)$	+	0	Unstable point	

TABLE 7: Assignment of each parameter in Case 3.

Parameter	$\alpha_1$	$\alpha_2$	$N$	$C_1$	$C_2$	$S_1$	$F_2$	$B_1$	$L_2$	$D_1$	$D_2$
Assignment	0.5	0.5	20	22	24	10	11	15	32	10	20

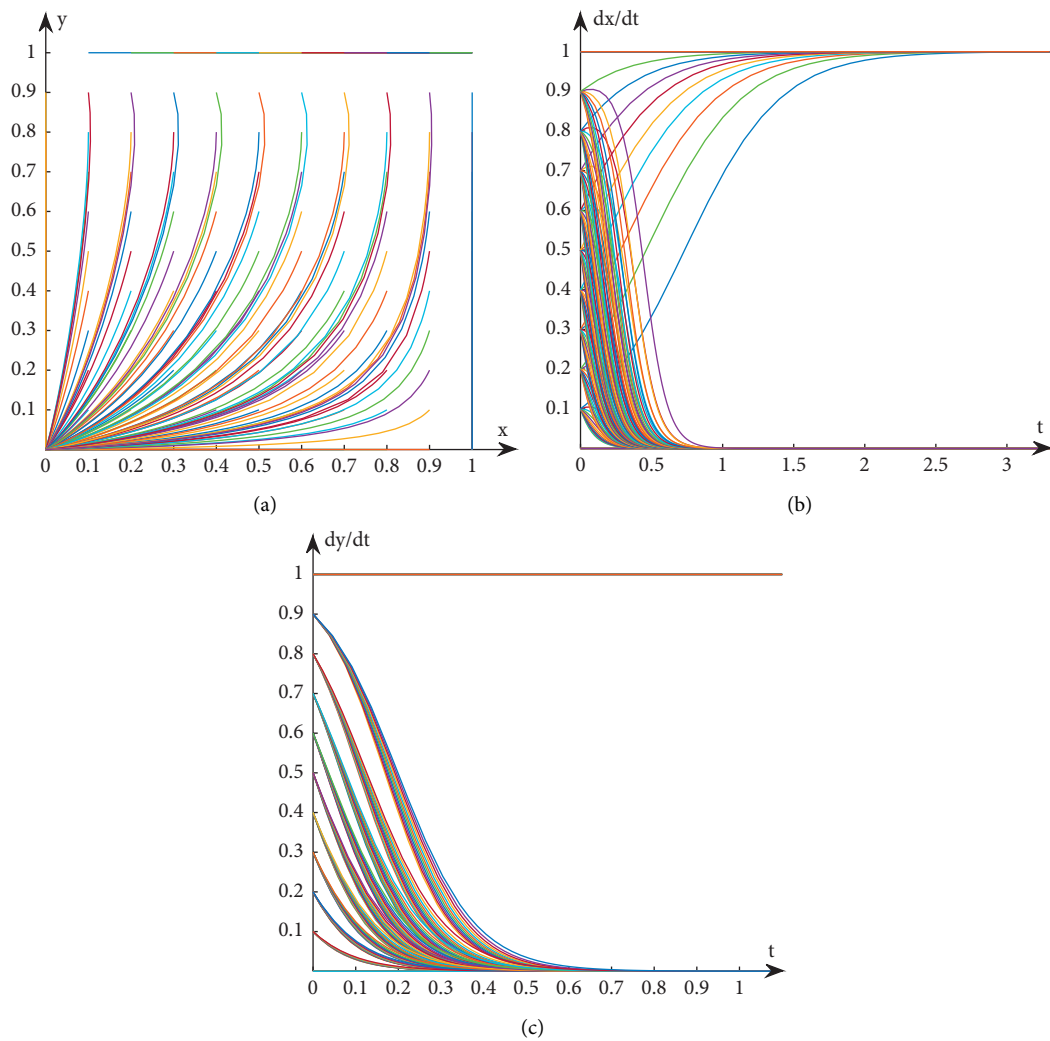


FIGURE 3: Evolution path diagram in Case 3. (a) Evolution path of system. (b) Evolution path of book publishers. (c) Evolution path of book retailers.

participate in the optimization, so (0,0) is an evolutionary stability strategy.

- (4) Case 4:  $\beta_1 < 0, \beta_2 < 0$ . In the same way, this paper obtains the expression of the determinant and trace of the equilibrium point of the system, as shown in Table 8.

Similarly, the parameters are assigned, as shown in Table 9. This paper can get the dynamic evolution path and trend of replication of book publishers and book retailers in Case 4, as shown in Figure 4.

It can be seen from Case 4 that when  $\beta_1 < 0, \beta_2 < 0$ , that is,  $B_1 - C_1 + N - \alpha_1 N < D_1 - S_1, L_2 - C_2 - \alpha_2 N < D_2 - F_2$ , the net revenue of book publishers and book retailers participating in the optimization is less than the net revenue of not participating in optimization. Thus, both parties chose not to participate in the optimization strategy.

## 4. Sensitivity Analysis of Group Behavior Evolution System

*4.1. The Influence of the Initial State on the Evolution Path of the System.* By changing the initial state of the proportion of book publishers participating in optimization  $X$  and the proportion of book retailers participating in optimization  $Y$ , this paper further analyzed the influence of different proportions of initial states on the evolution path and stability strategy of the system under the condition of Case 1 and obtained the simulation results by using MATLAB software.

### 4.1.1. The Evolution Path of Book Publishers

- (1) Discuss the evolution path of book publishers under different proportions of book retailers' selection strategies. It is assumed that when the proportion of book retailers participating in the optimization is selected with a probability of 0.5, the book publishers participate with the probability of 0.1, 0.3, 0.5, 0.7, and 0.9, respectively. The simulation results are shown in Figure 5.
- (2) Discuss the evolution path of book publishers under different proportions of book publishers' selection strategies. It is assumed that when the proportion of book retailers participating in the optimization is selected with a probability of 0.5, the book publishers participate with the probability of 0.1, 0.3, 0.5, 0.7, and 0.9, respectively. The simulation results are shown in Figure 6.

### 4.1.2. The Evolutionary Path of Book Retailers

- (1) Discuss the evolution path of book retailers under different proportions of book publishers' selection strategies. It is assumed that when the proportion of book retailers participating in the optimization is selected with a probability of 0.5, the book publishers participate with the probability of 0.1, 0.3, 0.5, 0.7,

and 0.9, respectively. The simulation results are shown in Figure 7.

- (2) Discuss the evolution path of book retailers under different proportions of book retailers' selection strategies. It is assumed that when the proportion of book publishers participating in the optimization is selected with a probability of 0.5, the book retailers participate with the probability of 0.1, 0.3, 0.5, 0.7, and 0.9, respectively. The simulation results are shown in Figure 8.

As can be seen from Figures 5–8, when the initial proportion of book publishers or book retailers choosing to participate in the optimization is relatively low, for example, when the value is 0.1, 0.3, or 0.5, its evolutionary path tends to the unsatisfactory state of not participating in the optimization. When the initial proportion of book publishers or book retailers choosing to participate in the optimization is relatively high, such as 0.7 or 0.9, its evolutionary path tends to the ideal state of participating in the optimization.

### 4.2. The Influence of Parameter Value Changes on the Evolutionary Game System

*4.2.1. Theoretical Analysis.* When the net revenue of book publishers and book retailers participating in the optimization is greater than the net revenue of not participating in the optimization, that is,  $\beta_1 > 0, \beta_2 > 0$ , the evolutionary system will eventually tend to the (0, 0) strategy that neither side will choose to participate in the optimization or both sides tend to choose (1, 1) strategy that participates in the optimization. Therefore, (0, 0) and (1, 1) are evolutionary stable strategies of the system, and the point  $D(X_0, Y_0)$  is the saddle point. The specific evolution process is shown in Figure 9.

The regional area of quadrilateral AOCD represents the probability that neither book publishers nor book retailers will participate in the optimization, while the regional area of quadrilateral ABCD represents the probability that both book publishers and book retailers will participate in the optimization. Therefore, the area size of the quadrilateral can represent the possibility that the two sides of the game choose different corresponding strategies, and its size is related to the position of the saddle point  $(X_0, Y_0)$ . The probability of system stability strategy can be changed by changing the numerical size of the saddle point parameter to guide the evolutionary game system to evolve towards the ideal state, where  $X_0 = (C_2 - F_2)/(L_2 - D_2 - \alpha_2 N)$ ,  $Y_0 = (C_1 - S_1)/(B_1 - D_1 + N - \alpha_1 N)$ .

Let  $S_{ABCD}$  represent the area of quadrilateral ABCD and  $S_{AOCD}$  represent the area of quadrilateral AOCD.

$$S_{ABCD} = 1 - \frac{1}{2} \left( \frac{C_2 - F_2}{L_2 - D_2 - \alpha_2 N} + \frac{C_1 - S_1}{B_1 - D_1 + N - \alpha_1 N} \right), \quad (8)$$

$$S_{AOCD} = \frac{1}{2} \left( \frac{C_2 - F_2}{L_2 - D_2 - \alpha_2 N} + \frac{C_1 - S_1}{B_1 - D_1 + N - \alpha_1 N} \right). \quad (9)$$

TABLE 8: The local stability of the equilibrium point in Case 4.

Equilibrium point	det ( $J$ )	tr ( $J$ )	Local stability	Strategy profile
(0, 0)	+	-	ESS	(Non-participation, non-participation)
(0, 1)	-	Uncertain	Saddle point	(Participation, non-participation)
(1, 0)	-	Uncertain	Saddle point	(Non-participation, participation)
(1, 1)	+	+	Unstable point	(Participation, participation)
$(X_0, Y_0)$	-	0	Saddle point	

TABLE 9: Assignment of each parameter in Case 4.

Parameter	$\alpha_1$	$\alpha_2$	$N$	$C_1$	$C_2$	$S_1$	$F_2$	$B_1$	$L_2$	$D_1$	$D_2$
Assignment	0.5	0.5	20	22	24	4	11	15	32	10	20

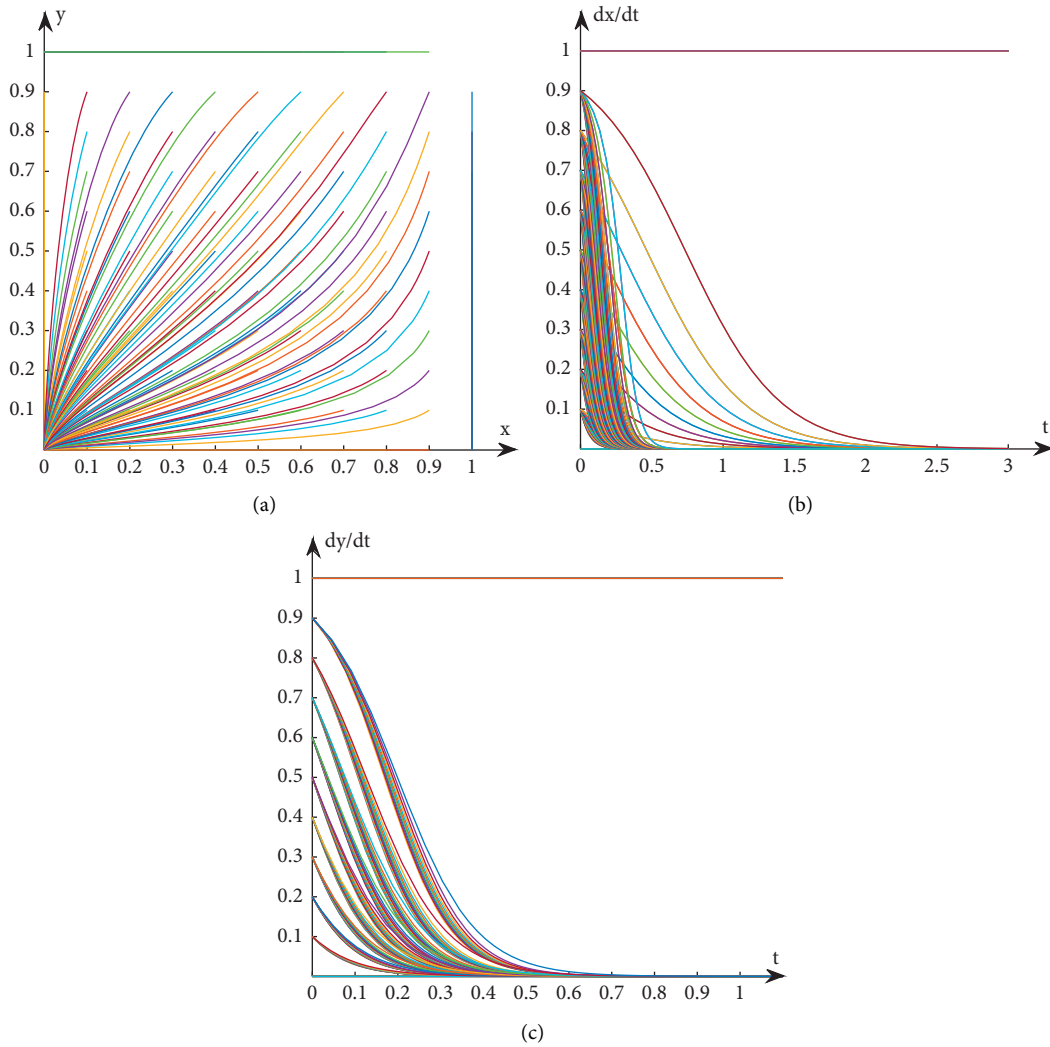


FIGURE 4: Evolution path diagram in Case 4. (a) Evolution path of system. (b) Evolution path of book publishers. (c) Evolution path of book retailers.

According to formulas (8) and (9), the parameters in the game model of book publishers-book retailers are analyzed, and Table 10 is obtained.

From Table 10, it can be seen that when the production and sales risk ratio of book publishers decreases and the cost

of reverse logistics decreases, the cost of book publishers and book retailers participating in optimization reduces and the revenue increases, and the cost of not participating in optimization reduces and the revenue decreases, and the system will reach the (1, 1) state more quickly. When  $N$

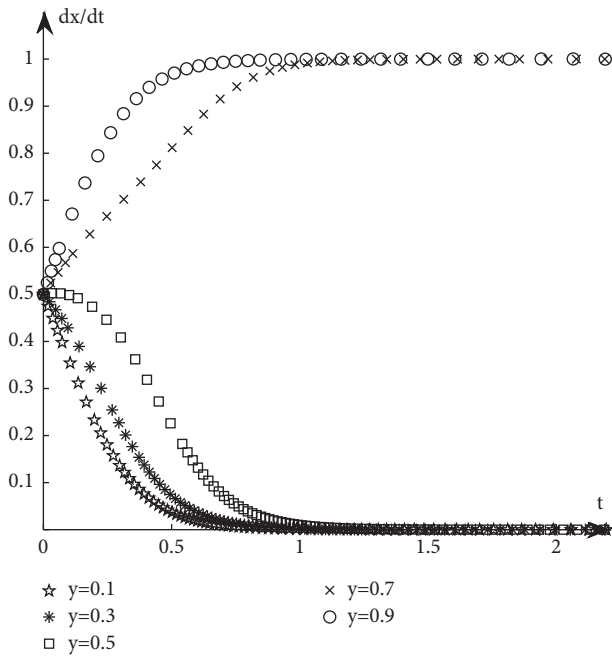


FIGURE 5: Evolution path when the initial state of  $X$  is the same.

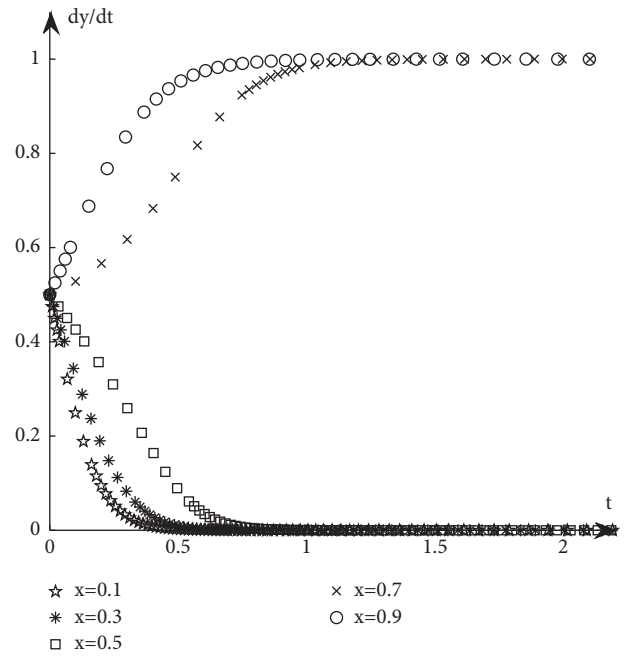


FIGURE 7: Evolution path when the initial state of  $X$  is the same.

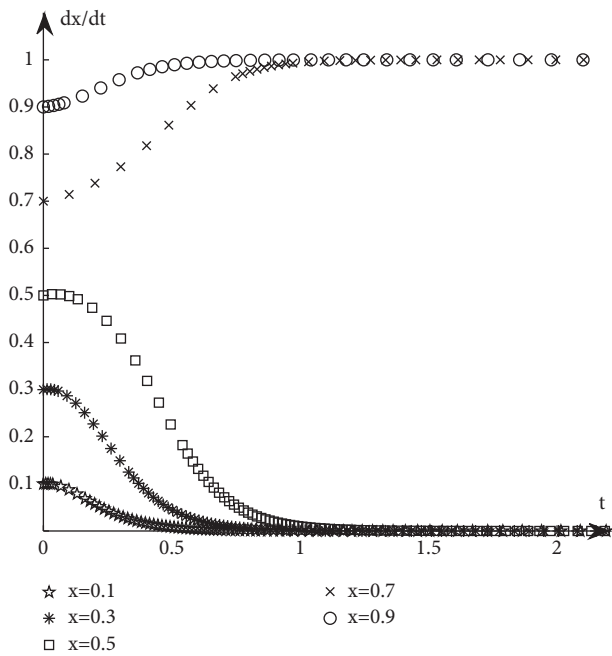


FIGURE 6: Evolution path when the initial state of  $Y$  is the same.

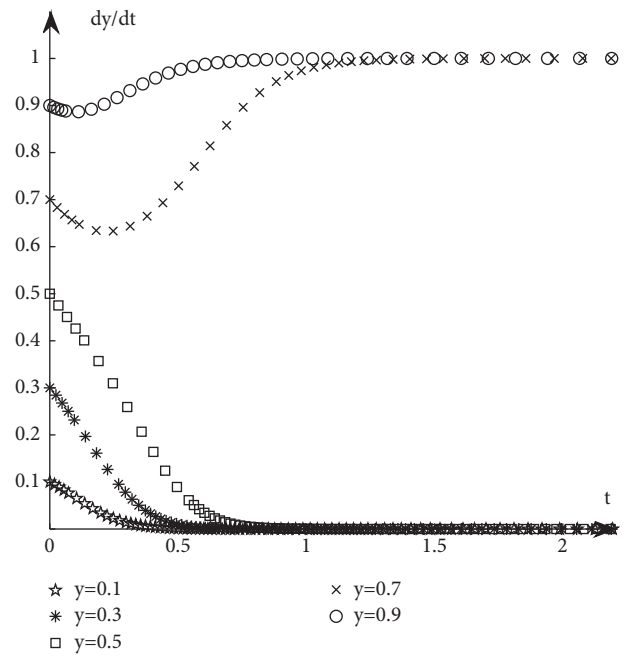


FIGURE 8: Evolution path when the initial state of  $Y$  is the same.

reaches condition  $X=1$ , the probability of both companies participating in the optimization of book reverse logistics is the greatest.

**4.2.2. Numerical Simulation.** To study the influence of parameter values on the evolution path and trend of the system, this paper selects an appropriate initial state,  $X=0.8, Y=0.8$ , and changes the values of  $\alpha_1, \alpha_2, N, B_1, L_2, C_1, C_2, S_1, F_2, D_1$ , and  $D_2$  at the same time. The specific

values are shown in Table 11. Group A is the control group, and the other groups are the experimental groups.

- (1) The influence of  $N$  changes on the evolution path. Keep other parameters unchanged and change the value of  $N$  (as shown in  $E$  in Table 11) to obtain a simulation diagram of the evolution trend of book publishers and book retailers, as shown in Figure 10.

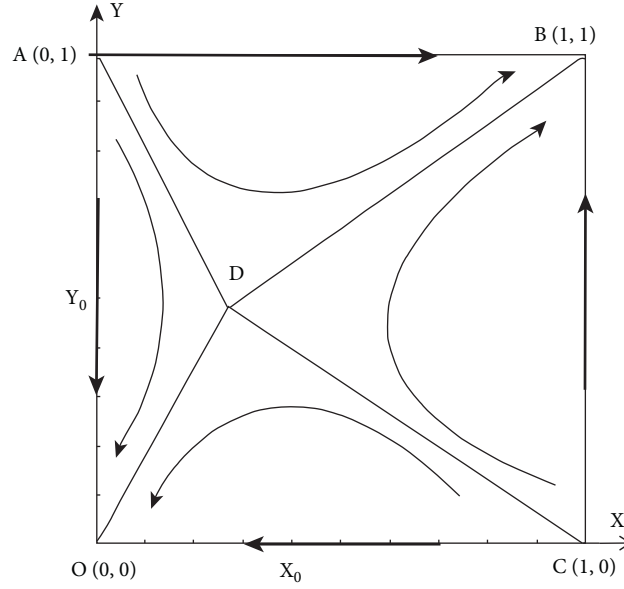


FIGURE 9: Phase diagram of system evolution in Case 1.

TABLE 10: Influence of parameter changes on evolution stability results.

Variation of parameters	Basis of judgment	Result
$\alpha_1 \uparrow$	$\partial S_{ABCD}/\partial \alpha_1 = -(N(C_1 - S_1))/(2(B_1 - D_1 + N - \alpha_1 N)^2) < 0,$ $\partial S_{A OCD}/\partial \alpha_1 = (N(C_1 - S_1))/(2(B_1 - D_1 + N - \alpha_1 N)^2) > 0$	Both parties prefer not to participate in the optimization, and the evolutionary game system converges to the point (0, 0).
$\alpha_2 \uparrow$	$\partial S_{ABCD}/\partial \alpha_2 = -(N(C_2 - F_2))/(2(L_2 - D_2 - \alpha_2 N)^2) < 0,$ $\partial S_{A OCD}/\partial \alpha_2 = (N(C_2 - F_2))/(2(L_2 - D_2 - \alpha_2 N)^2) > 0$	
$D_1 \uparrow$	$\partial S_{ABCD}/\partial D_1 = -(C_1 - S_1)/(2(B_1 - D_1 + N - \alpha_1 N)^2) < 0,$ $\partial S_{A OCD}/\partial D_1 = (C_1 - S_1)/(2(B_1 - D_1 + N - \alpha_1 N)^2) > 0$	
$D_2 \uparrow$	$\partial S_{ABCD}/\partial D_2 = -(C_2 - F_2)/(2(L_2 - D_2 - \alpha_2 N)^2) < 0,$ $\partial S_{A OCD}/\partial D_2 = (C_2 - F_2)/(2(L_2 - D_2 - \alpha_2 N)^2) > 0$	
$C_1 \uparrow$	$\partial S_{ABCD}/\partial C_1 = -C_1/(2(B_1 - D_1 + N - \alpha_1 N)) < 0,$ $\partial S_{A OCD}/\partial C_1 = C_1/(2(B_1 - D_1 + N - \alpha_1 N)) > 0$	Both parties are more inclined to participate in the optimization, and the evolutionary game system converges to the point (1, 1).
$C_2 \uparrow$	$\partial S_{ABCD}/\partial C_2 = -C_2/(2(L_2 - D_2 - \alpha_2 N)) < 0,$ $\partial S_{A OCD}/\partial C_2 = C_2/(2(L_2 - D_2 - \alpha_2 N)) > 0$	
$S_1 \uparrow$	$\partial S_{ABCD}/\partial S_1 = S_1/2(B_1 - D_1 + N - \alpha_1 N) > 0,$ $\partial S_{A OCD}/\partial S_1 = -S_1/(2(B_1 - D_1 + N - \alpha_1 N)) < 0$	
$F_2 \uparrow$	$\partial S_{ABCD}/\partial F_2 = F_2/(2(L_2 - D_2 - \alpha_2 N)) > 0,$ $\partial S_{A OCD}/\partial F_2 = -F_2/(2(L_2 - D_2 - \alpha_2 N)) < 0$	
$B_1 \uparrow$	$\partial S_{ABCD}/\partial B_1 = C_1 - S_1/(2(B_1 - D_1 + N - \alpha_1 N)^2) > 0,$ $\partial S_{A OCD}/\partial B_1 = -C_1 - S_1/(2(B_1 - D_1 + N - \alpha_1 N)^2) < 0$	Both book publishers and book retailers have the greatest probability of participating in the optimization.
$L_2 \uparrow$	$\partial S_{ABCD}/\partial L_2 = (C_2 - F_2)/(2(L_2 - D_2 - \alpha_2 N)^2) > 0,$ $\partial S_{A OCD}/\partial L_2 = -(C_2 - F_2)/(2(L_2 - D_2 - \alpha_2 N)^2) < 0$	
$N$	$\text{当 } C_2 - F_2/(L_2 - D_2 - \alpha_2 N)^2 = C_2 - F_2/(B_1 - D_1 + N - \alpha_1 N)^2 \text{ 时}$	

It can be seen from Figure 10 that when  $N$  increases, the proportion of book retailers participating in the optimization greatly decreases, which means that the willingness of book retailers to participate in the optimization decreases, and the convergence speed of the system slows down; the proportion of book publishers participating in the optimization has a small increase, but its impact on the evolution path is too small, so the system will more slowly reach the

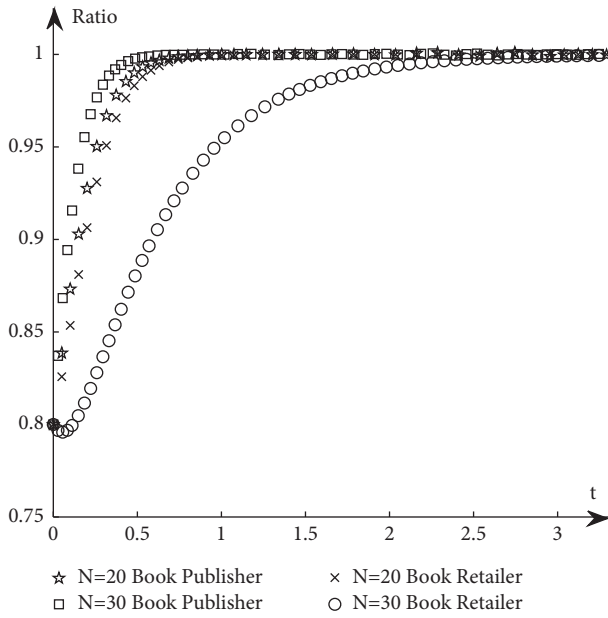
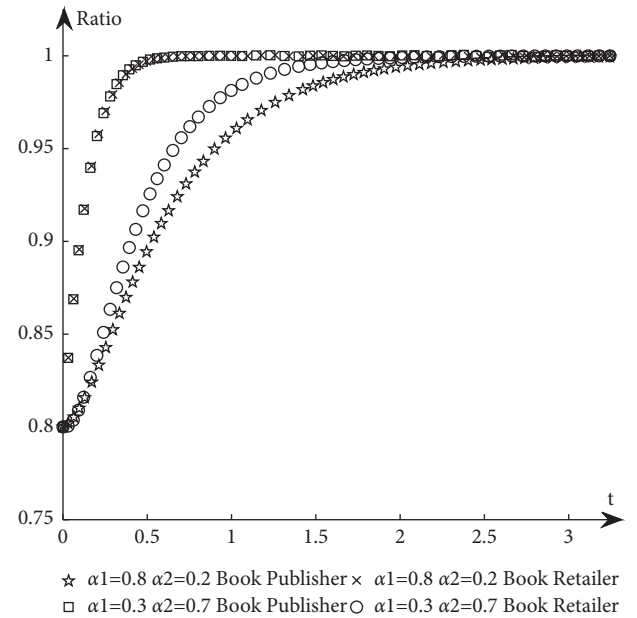
ideal state where both parties participate in the optimization.

- (2) The influence of  $\alpha_1$  and  $\alpha_2$  changes on the evolution path. Change the size of  $\alpha_1$  and  $\alpha_2$ , keep  $\alpha_1 + \alpha_2 = 1$  and other parameters unchanged, study their influence on the evolutionary system (as shown in  $G_1$  and  $G_2$  in Table 11), and obtain the evolutionary trend simulation diagram of book publishers and book retailers, as shown in Figure 11.



TABLE 11: Parameter values.

Parameter	$\alpha_1$	$\alpha_2$	$N$	$C_1$	$C_2$	$S_1$	$F_2$	$B_1$	$L_2$	$D_1$	$D_2$
$A$	0.5	0.5	20	17	24	10	11	15	50	10	20
$E$	0.5	0.5	30	17	24	10	11	15	50	10	20
$G_1$	0.8	0.2	20	17	24	10	11	15	50	10	20
$G_2$	0.3	0.7	20	17	24	10	11	15	50	10	20
$H_1$	0.5	0.5	20	22	24	10	11	15	50	10	20
$H_2$	0.5	0.5	20	17	19	10	11	15	50	10	20
$J_1$	0.5	0.5	20	17	24	5	11	15	50	10	20
$J_2$	0.5	0.5	20	17	24	10	16	15	50	10	20
$K_1$	0.5	0.5	20	17	24	10	11	25	50	10	20
$K_2$	0.5	0.5	20	17	24	10	11	15	60	10	20
$P_1$	0.5	0.5	20	17	24	10	11	15	50	5	20
$P_2$	0.5	0.5	20	17	24	10	11	15	50	10	15

FIGURE 10: The simulation diagram of the influence of  $N$  on the evolution path.FIGURE 11: The simulation diagram of the influence of  $\alpha_1$  and  $\alpha_2$  on the evolution path.

It can be seen from Figure 11 that when the production and sales risk ratio of book publishers decreases and the production and sales risk ratio of book retailers increases, it will speed up book publishers to reach the ideal state of participating in the optimization and slow down book retailers that have reached the ideal state of participating in the optimization. Under the condition of the same production and sales risk ratio, the proportion of book publishers participating in the optimization is greater than that of book retailers, indicating that the possibility of book publishers participating in the optimization is greater than the possibility of book retailers participating in the optimization when the production and sales risks are relatively high.

- (3) The influence of  $C_1$  and  $C_2$  changes on the evolution path. Keep other parameters unchanged and change the values of  $C_1$  and  $C_2$ , respectively (as shown in  $H_1$  and  $H_2$  in Table 11), to obtain a simulation diagram

of the evolution trend of book publishers and book retailers, as shown in Figure 12.

It can be seen from Figure 12 that when the cost of book publishers and book retailers participating in the optimization increases, the revenue of participating in the optimization will decrease, and the willingness to participate in the optimization will decrease, slowing down the speed at which the system converges to the ideal state participating in the optimization. On the contrary, it will increase the revenue of book publishers and book retailers participating in the optimization, increase the willingness to participate in the optimization, and accelerate the speed of convergence.

- (4) The influence of  $S_1$  and  $F_2$  changes on the evolution path. Keep other parameters unchanged and change the values of  $S_1$  and  $F_2$ , respectively (as shown in  $J_1$  and  $J_2$  in Table 11), to obtain a simulation diagram of

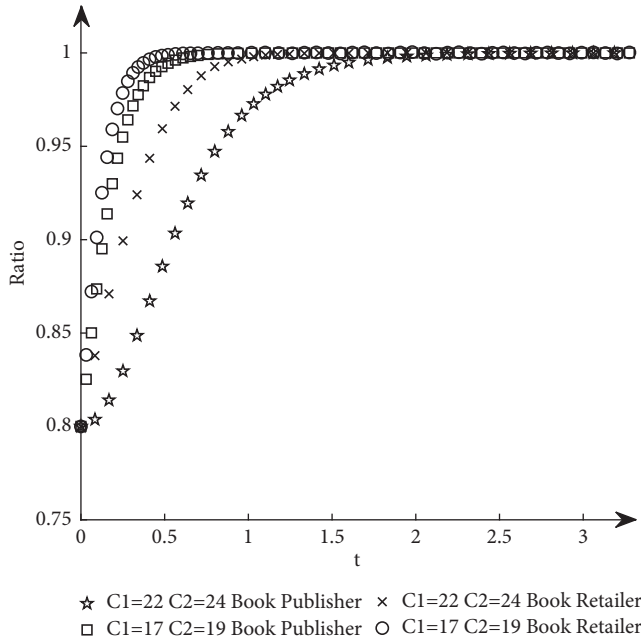


FIGURE 12: The simulation diagram of the influence of  $C_1$  and  $C_2$  on the evolution path.

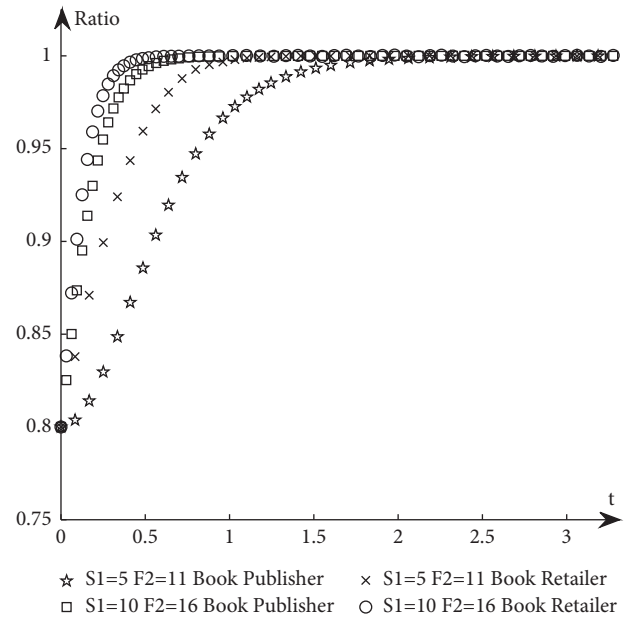


FIGURE 13: The simulation diagram of the influence of  $S_1$  and  $F_2$  on the evolution path.

the evolution trend of book publishers and book retailers, as shown in Figure 13.

As can be seen from Figure 13, when  $S_1$  and  $F_2$  increase, the revenue of book publishers and book retailers not participating in the optimization will decrease, and both sides' willingness to participate in the optimization will be stronger, accelerating the convergence speed of the system and reaching the ideal state of participating in optimization more quickly. On the contrary, the result is the opposite.

- (5) The influence of  $B_1$  and  $L_2$  changes on the evolutionary path. Keep other parameters unchanged and change the values of  $B_1$  and  $L_2$  respectively (as shown in  $K_1$  and  $K_2$  in Table 11), to obtain a simulation diagram of the evolution trend of book publishers and book retailers, as shown in Figure 14.

It can be seen from Figure 14 that when the revenue of book publishers and book retailers participating in the optimization increases, the willingness to participate in the optimization will increase, speeding up the convergence of the system. Conversely, the willingness of book publishers and book retailers to participate in the optimization will decrease, which will slow down the system's convergence speed.

- (6) The influence of  $D_1$  and  $D_2$  changes on the evolutionary path. Keep other parameters unchanged and change the values of  $D_1$  and  $D_2$ , respectively (as shown in  $P_1$  and  $P_2$  in Table 11), to obtain a simulation diagram of the evolution trend of book publishers and book retailers, as shown in Figure 15.

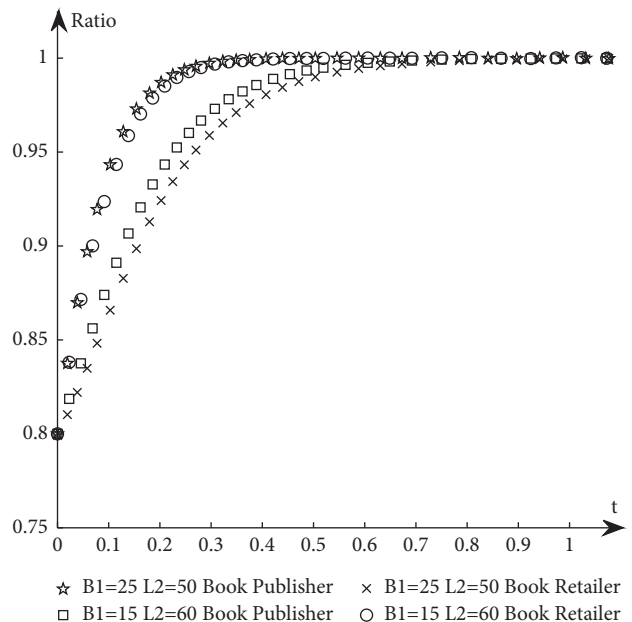


FIGURE 14: The simulation diagram of the influence of  $B_1$  and  $L_2$  on the evolution path.

It can be seen from Figure 15 that when book publishers and book retailers do not participate in the optimization, revenue increases, and book publishers and book retailers will reduce their willingness to participate in the optimization, slowing down system convergence speed. Conversely, the willingness of book publishers and book retailers

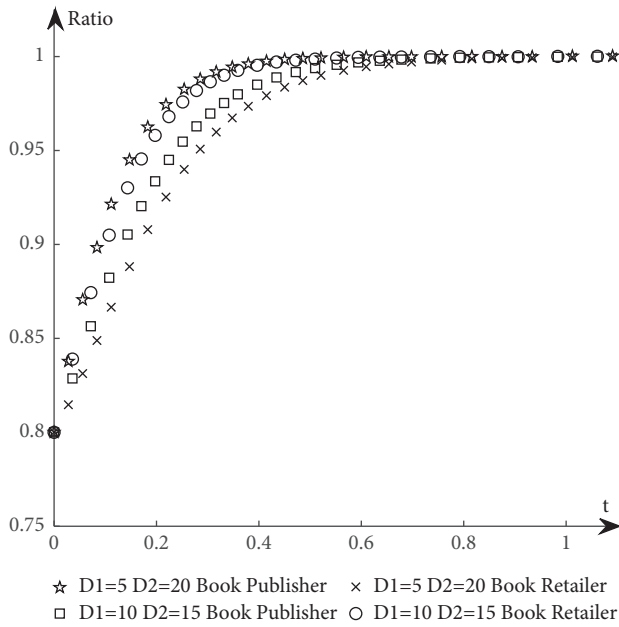


FIGURE 15: The simulation diagram of the influence of  $D_1$  and  $D_2$  on the evolution path.

to participate in the optimization will increase, speeding up system convergence.

## 5. Conclusion

Based on the research of reverse logistics and evolutionary game theory by domestic and foreign scholars, this paper constructs an evolutionary game model between book publishers and book retailers, obtains two evolutionary stable strategies (participation, participation) and (non-participation, non-participation), and uses MATLAB to carry out numerical simulation to analyze the evolution path of the two companies under different evolutionary and stable strategies. The main conclusions are as follows. (1) Through exploring the stability of the interest groups of the evolutionary game in different situations and the influence of the initial state on the evolution path of the system, this paper finds that only when the net revenue of both parties participating in optimization is greater than the net revenue of not participating in optimization, the evolution game system can reach the ideal state where both parties participate in the optimization. When the initial ratio of book publishers and book retailers participating in the optimization is greater than or equal to 0.7, the system will eventually evolve into the ideal state of participating in optimization. (2) Through analyzing the influence of parameter value changes on the evolutionary game system, this paper found that when the cost of participating in the book reverse logistics reaches the condition of  $(C_2 - F_2)/(L_2 - D_2 - \alpha_2 N)^2 = (C_1 - S_1)/(B_1 - D_1 + N - \alpha_1 N)^2$ , the probability of both parties participating in the optimization is the greatest. Under the condition of the same production and sales risk ratio, the proportion of book publishers participating in the optimization is greater than the proportion of book retailers

participating in the optimization, indicating that when the production and sales risk ratio is relatively high, the possibility of book publishers participating in the optimization is greater than the possibility of book retailers participating in optimization. When the production and sales risk ratio of book publishers decreases and the cost of reverse logistics decreases, the cost of book publishers and book retailers participating in optimization reduces and the revenue increases, and the cost of not participating in optimization reduces and the revenue decreases, and the system will more quickly reach the ideal state where both parties participate in the optimization.

## Data Availability

The data for all Figures used to support the findings of this study are included within the article.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported in part by Scientific Research Program of Beijing Education Commission under Grant KM201910015004.

## References

- [1] J. Pu, "Forty years of book inventory changes-reflections on cultural production," *Research on Transmission Competence*, vol. 3, pp. 10–12+35, 2019.
- [2] C. H. Zhang, "Research on book reverse logistics based on closed-loop supply chain," *China Publishing Journal*, vol. 20, pp. 56–58, 2010.
- [3] W. Ma, "Discussion on the optimization of book centralized inventory based on reverse logistics," *Logistics Technology*, vol. 33, pp. 303–305, 2014.
- [4] Z. Ning, "Research on reverse logistics of book publishing industry based on green supply chain," *Logistics Technology*, vol. 33, pp. 248–250, 2014.
- [5] T. R. Morgan, R. G. Richey, and C. Defee, "Resource commitment and sustainability: a reverse logistics performance process model," *International Journal of Physical Distribution and Logistics Management*, vol. 48, no. 2, pp. 164–182, 2018.
- [6] A. Faustino, P. Cortés-Pellicer, D. Pérez-Perales, and A. Mengual-Recuerda, "A reference model of reverse logistics process for improving sustainability in the supply chain," *Sustainability*, vol. 13, no. 18, Article ID 10383, 2021.
- [7] J. Yi, "Stakelberg game analysis of reverse supply chain with group evolutionary characteristic," *Journal of Shenzhen University Science and Engineering*, vol. 30, no. 1, pp. 103–108, 2013.
- [8] W. Xia, Q. Xia, N. Zhou, and D. Zhou, "Based on reverse logistics of the game model of the manufacturer recycles waste packaging," *Journal of Chongqing University of Technology (Natural Science)*, vol. 32, pp. 135–140, 2018.
- [9] J. Ye, H. Ma, and X. Li, "Evolutionary game analysis of the supervision of the garbage classification and recycling reverse logistics," *Logistics Sci-Tech*, vol. 38, pp. 78–82, 2015.

- [10] Y. Chen, "The management game analysis of the third-party recyclers in closed-loop supply chain," *Nanjing University*, Nanjing, China, 2015.
- [11] T. Wang and G. Nan, "Decision making model based on game theory for retailer recycling mode," *Journal of Systems Engineering*, vol. 30, pp. 790–803, 2015.
- [12] D. Chen, J. Ignatius, D. Sun et al., "Reverse logistics pricing strategy for a green supply chain: a view of customers' environmental awareness," *International Journal of Production Economics*, vol. 217, pp. 197–210, 2019.
- [13] Y. Shi, J. Wang, and Z. Zhang, "Pricing strategy based on game theory for two-stage reverse supply chain," *Industrial Engineering Journal*, vol. 23, pp. 45–50+58, 2020.
- [14] Y. Ding, D. Feng, S. Jin, Ye Zhang, and C. Wang, "Pricing decision for remanufacturing reverse logistics based on game theory," *Control and Decision*, vol. 33, pp. 749–758, 2018.
- [15] Y. Li and F. Wang, "The game analysis of steel enterprise implementating of scrap reverse logistics under low carbon economy," *Journal of Anhui University of Technology*, vol. 34, pp. 29–34, 2017.

## *Retraction*

# **Retracted: Innovative Research on Urban Community Governance Decision-Making Relying on Distributed High-Performance Computing Blockchain Key Algorithms**

### **Computational Intelligence and Neuroscience**

Received 3 October 2023; Accepted 3 October 2023; Published 4 October 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### **References**

- [1] J. Wen and P. Xie, "Innovative Research on Urban Community Governance Decision-Making Relying on Distributed High-Performance Computing Blockchain Key Algorithms," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4078014, 13 pages, 2022.

## Research Article

# Innovative Research on Urban Community Governance Decision-Making Relying on Distributed High-Performance Computing Blockchain Key Algorithms

Jun Wen <sup>1</sup> and Peihong Xie<sup>2</sup>

<sup>1</sup>School of Cultural Creative Industries Management, Shanghai Institute of Visual Arts, Shanghai 201620, China

<sup>2</sup>School of Management, Shanghai University of International Business and Economics, Shanghai 201620, China

Correspondence should be addressed to Jun Wen; wenjun@siva.edu.cn

Received 18 January 2022; Accepted 16 February 2022; Published 15 March 2022

Academic Editor: Daqing Gong

Copyright © 2022 Jun Wen and Peihong Xie. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Since the country began to go global, the country's economy has developed rapidly, cultural exchanges between countries have become more and more frequent, and foreign cultures have begun to gradually spread to the country. At the same time, through the absorption of foreign community governance experience, domestic research on community governance has also begun to be put on the agenda. This paper aims to study the innovative exploration of urban community governance mechanisms through key high-performance computing algorithms. To this end, this paper proposes a combination of the Bayesian algorithm and distributed high-performance computing to analyze and explore the governance and management methods of the community through its efficient and stable computing power, and derive the most suitable community governance mechanism. An experiment was also set up for comparative analysis. The experimental results show that the community governance mechanism derived from the key distributed high-performance computing algorithm improves the community governance capability by 19.4%, effectively improving the community governance and management issues.

## 1. Introduction

In recent years, with the rapid development of Chinese politics, economy, and society, the process of urbanization has accelerated, and residential privatization has continued to advance. In a sense, Chinese cities are divided into multiple functional communities. Community is the basic unit of social management and bears the responsibility of building a harmonious society. At the beginning of 2010, the outline of China's national economic and social development's 12-year plan proposed to fully implement the construction of urban communities, improve the new type of community management and service system, and build communities. It promotes the urbanization process of community construction, expands the management model, improves the service management system, accelerates the pace of marketization and modernization, and builds a harmonious socialist society.

The themes of construction, development, and governance of urban communities have received more and more praise from political parties, experts at all levels, scholars, and urban residents. Moreover, with the further development of urban community construction, the diversity governance of urban communities has also produced huge social benefits. Urban community governance is an important foundation of urban governance in the new era, the basic work of modern city construction, and an important part of the reform of modern city management systems. According to specific conventions, it is recommended to build an urban community governance mechanism. By strengthening the construction of adjusted governance mechanisms such as participation in decision-making, response governance, and public consultation, the opposition in urban community governance in the policy system will be fundamentally resolved and alleviated. In order to



comprehensively improve the efficiency of urban community governance, it is necessary to promote the improvement of Chinese urban governance system and the modernization of governance capabilities. Through the use of distributed computing and key blockchain algorithms, it is very necessary to govern urban communities.

The research on the construction of foreign cities is relatively mature compared with China. Research uses the path from practice to theory, and then applies theory to practice. The research results mainly focus on the impact of social changes in urban communities, the research on the operating mechanism and process of urban communities, and the types of urban communities. A new parallel approximate Bayesian computational sequential Monte Carlo (pABC SMC) algorithm proposed by Jagiella et al. allows robust data-driven modeling of multi-scale biological systems, and proves the feasibility of multi-scale model parameterization through statistical inference [1]. However, the algorithm is mainly based on the parameters obtained by inference on the theoretical basis, and has not been verified by the experimental area. The interactive high-performance computing proposed by Mundani et al. is undoubtedly beneficial to many computing science and engineering applications, as long as the simulation results be visualized in real time during the calculation process. Because a few gigabytes of data per second is not uncommon for running simulations, a new method based on the sliding window technology is introduced and allows users to interactively study the large-scale and small-scale effects of the simulation results [2]. The problem of real-time visualization processing in the calculation process has not been solved well. JUDEHOWELL has studied the process of adaptive governance and authoritarian resilience from the perspective of community governance, bringing together unrelated research fields so far. He believes that adaptation will have unintended consequences and may threaten stability, but it will not necessarily lead to a political crisis [3]. Mullen has spent decades developing tools for the High-Performance Computing (HPC) community, teaching practitioners to use the power of parallel and distributed computing to familiarize student users with their target systems. The course materials are presented through the lens of common HPC use cases and parallelization strategies so that students are able to achieve performance improvements on their HPC systems [4]. However, the learning path of the students is still not clear enough, and the relevant display is not enough. Palmer B introduced the GridPACK™ framework, which is designed to help grid engineers develop modeling software that can run on high-performance computers. GridPACK™ contains modules for setting up a distributed grid network, assigning buses and branches with arbitrary behavior to the network, creating distributed matrices and vectors, and using parallel linear and nonlinear solvers to solve algebraic equations [5]. But there is no clear introduction to how the framework works on related modeling software. The problem of Turek simulating micro-scale urban traffic in a large environment provides a huge opportunity for the use of HPC systems. The parallel implementation of this

calculation is not simple. The simulation proposed in his article is based on the concept of controlled desynchronization of computation, which does not violate the model. The implementation of the Erlang language uses the Erlang distribution mechanism to build and manage computing clusters [6]. The difficulty in the parallel realization of the calculation process is still lack of explanation, which leads to the lack of clear understanding of the HPC system. Alawneh L believes that if you use trace analysis, you can more easily understand the interaction between the processes of the high-performance computing (HPC) system. So, he proposed a novel tracking abstraction method designed to facilitate the analysis of large-scale execution tracking generated by HPC applications. Their method allows the automatic segmentation of large traces into smaller and meaningful clusters, which reflect the various execution stages of the trace scenario [7, 8]. But there is no good elaboration on the principle of this novel tracking abstract method. The purpose of Arya and Wijaya's research is to let people know more about the role of women on the coast of Surabaya. The role of coastal women is not limited to their family role, and through social capital management, coastal women can also contribute to local governance. Research has found that coastal women groups can achieve success in social bonds, bridges, and connections. Under the leadership of a sufficiently effective women's group to manage the organization, people realized the need for a savings movement and had a process of sharing information and knowledge. The social bridge capital shown for establishing connections with other groups also provides a foundation for the construction of an urban society from another angle, which is convenient for later development [9]. The above documents are still very detailed for their main technologies, and the research on some technologies is also very in-depth. However, the above-mentioned literature mainly studies the related application fields of high-performance computing, and the research on its actual operability is not enough, and it is easy to have problems in the actual use process.

The innovation of this paper is to use the key algorithm extraction of the blockchain combined with distributed high-performance computing as the technical support, the urban community governance theory as the theoretical support, and the research and analysis of the urban community governance as the research object. And in the experiment and analysis part, related experiments are designed to explore the effect of its governance, to ensure that the improved governance mechanism can be effectively implemented in practice. In addition, two key algorithms were selected for comparative analysis. On this basis, some more mature urban community governance theories in foreign countries were used to construct a governance path, learn from its experience, select the essence, and remove the dross, which combines the governance methods of China's socialist society to deduct the governance of urban communities, and with the support of high-performance computing capabilities, it deduces a governance method that is more suitable for most communities and improves the governance mechanism of the community.

## 2. High-Performance Computing Methods

*2.1. Blockchain Cloud Computing Method.* Cloud computing is a paid service, and users need to purchase corresponding cloud computing services on demand under the premise of fully evaluating their own situation. Cloud computing implements the abstraction of resources such as data computing, data storage, network transmission, and hardware equipment through key technologies such as virtualization, distributed parallel processing, and the Internet, and then provides these resources to users dynamically on demand. This provides users with a more flexible, convenient, and economical way to dynamically obtain computing resources and storage resources according to their needs [10]. For users, instead of purchasing and maintaining huge and expensive infrastructure and incurring huge costs, it is obviously better to dynamically and expandably rent the necessary cloud computing resources and storage resources according to their needs. In this way, users can not only avoid early-stage investment risks, but also avoid the inability of resources to meet demand due to the increase in the company's scale, which is very attractive to small- and medium-sized enterprises, especially start-ups. Although cloud computing is considered to be the service model of the next generation Internet and cloud computing services have become more and more perfect after this year's rapid development and research, data security and privacy protection are still the main challenges that currently restrict the further development of the cloud computing field (location information protection, mobile phone data protection, authority protection, etc.). Unlike the traditional computing model where users fully control data calculation and storage locally, cloud computing requires the user's data and physical servers to be centrally managed by the cloud service provider, and the user only retains some control rights over the leased virtual machine [11]. The advantages of traditional computing models compared with cloud computing models are mainly focused on the security protection of data because they completely control local data and storage, thus having higher security. The cloud computing method application service of blockchain is shown in Figure 1:

There are three common deployment models for cloud computing: public cloud, private cloud, and hybrid cloud. Users can choose to use different deployment models according to the type of data to be processed and the requirements for security and data management [12].

**Public Cloud:** In the public cloud, service providers use the Internet to provide users with computing resources including applications and storage. Public clouds are operated by third-party providers, and applications and service requests from different users are collectively run and processed on cloud servers, storage systems, and networks. Public cloud has the advantages of low cost, no maintenance, high scalability, and high reliability [13].

**Private Cloud:** A private cloud is a cloud computing service dedicated to a single user, which can be regarded as a user's private data service center. In a private cloud, infrastructure and services are managed and maintained on a dedicated network, and software and hardware services only

serve a single user. Private clouds can also be operated and maintained by third-party service providers. Private clouds are often favored by users who are sensitive to data security, such as government agencies and financial institutions. Private cloud services have the advantages of more flexibility, higher security and reliability, and high open scalability.

**Hybrid Cloud:** Hybrid cloud combines the user's infrastructure or private cloud and public cloud to obtain the advantages of both at the same time. In a hybrid cloud, data and applications can be securely transferred between private and public clouds for greater flexibility and more deployment options. Users can use the public cloud to perform operations with high storage capacity and low security while using the private cloud to perform secure operations on sensitive businesses. Hybrid cloud combines the advantages of the previous two, with advantages such as controllability and flexibility.

Cloud computing implements the abstraction of data computing, data storage, network transmission, and hardware equipment through key technologies such as resource management, distributed parallel processing, and the Internet, and then provides these resources to users dynamically on demand. Taking into account the above characteristics, this article adopts the hybrid cloud deployment model for related deployment. This service process mainly involves three aspects of technology, including resource management, Internet technology, and distributed computing. The research on distributed computing is the focus of this article. As shown in Figure 2, it is the key technology involved in the cloud computing process [14].

Resource management technology mainly includes virtualization technology and data center management technology. Virtualization technology is one of the core key technologies of cloud computing, which can abstract the details of physical hardware and provide virtualized resources for advanced applications. Virtual machine services with scalable performance can be provided to users through virtualization technology, which further improves the utilization efficiency and high scalability of physical servers in the case of multiple users. Data center is the core of the cloud computing architecture. Data center management technology is one of the most basic core technologies of cloud computing, which directly affects the performance of upper cloud data storage and computing services. Cloud computing data centers include technologies such as network topology, large-scale data storage, resource management and scheduling, and they need to have features such as autonomy, economies of scale, and scalability. In cloud computing, Internet technology is used to complete the interaction between the users and cloud resources. Cloud computing only requires the user's data and physical servers to be centrally managed by the cloud service provider, and the user only retains some control over the leased virtual machine. At present, most cloud computing service providers provide users with web-based cloud interactive interfaces. The B/S architecture-based approach not only reduces the development cycle, but also further reduces the resource occupation of the client [15].

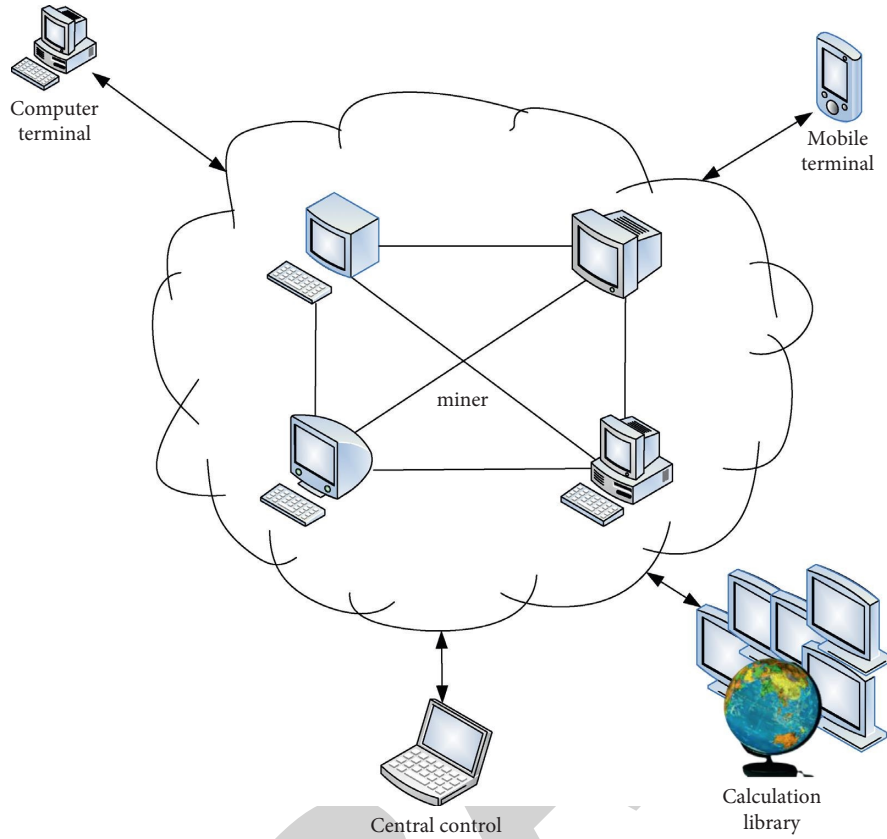


FIGURE 1: Application services of blockchain cloud computing.

## 2.2. SVM Algorithm Optimization for Distributed Computing

### 2.2.1. Classification of Classification Algorithms

(1) *Decision Tree Algorithm.* The decision tree method is to first find the most informative attributes in the sample database as the nodes of the decision tree, and then construct various branches of the nodes according to the attribute values. Then, repeat this process for each one. It is used to establish the branch of the next node and its branches [16].

(2) *Bayesian Classification.* Bayesian classification is a statistical classification method. The classification principle is to use the pre-probability of the object and use the Bayesian equation to calculate the post-probability. In other words, the object has the probability of belonging to a specific category, and the category with the largest posterior probability is selected as the class the object belongs to [17].

For a training sample, suppose its prior probability is  $P$ ; according to Bayesian classification, the probability obtained is

$$p(z) = p(z/si) \cdot p(si)/p(x). \quad (1)$$

If

$$p(si/z) = \text{Max}_j p(sj/Z), i = 1, 2, 3 \dots, N, \quad (2)$$

then there is  $x \in si$ ; substituting formula (1) into formula (2),

$$p(z/si)p(si) = \text{Max}_j [p(z/sj)p(sj)], i = 1, 2, 3, \dots, N, j = 1, 2, \dots, N. \quad (3)$$

This is the most commonly used Bayes classification decision benchmark. As a result of years of research, the Bayesian classification is widely used. However, because the Bayesian sample size is large enough, the sample properties must be independent of each other; it is difficult to achieve this in practical applications, and hence it is difficult to achieve theoretical results with this method [18].

(3) *Neural Network Classification.* As shown in Figure 3, the neural network classification is mainly composed of three layers to form a multilayer feedforward neural network (Input layer, hidden layer, output layer) [19]. The multilayer feedforward neural network consists of an input layer, one or more hidden layers, and an output layer. The Back Propagation [BP] algorithm learns on the multilayer feedforward neural network, and the [multilayer] feedforward neural network using the BP algorithm is called the BP neural network.

The most commonly used multilayer feedforward neural network lies in classification. By setting the number of sample attribute values and controlling the sample type of the output layer, various classifications can be performed. But the problem that it is easier to fall into the minimum is one of its inherent shortcomings.

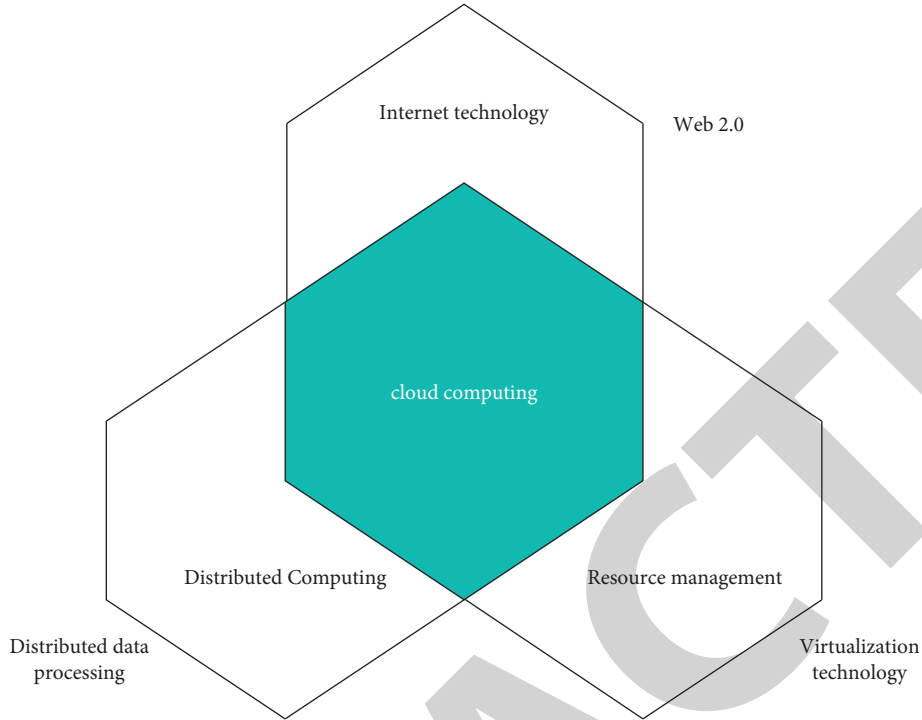


FIGURE 2: Key technologies involved in cloud computing.

(4) *Linear Classification.* The learning goal of a linear classifier is to find a classification hyperplane in the  $n$ -dimensional data space, and its equation can be expressed as

$$\omega^T z + b = 0. \quad (4)$$

As shown in Figure 4, in the two-dimensional plane, the solid line is the classification plane. We can make the classification function as

$$f(Z) = \omega^T z + b. \quad (5)$$

At the same time, use  $z$  to represent the “geometric distance” of the classification hyperplane;

$$\gamma = y \frac{\omega^T z + b}{\|\omega\|} = y \frac{f(z)}{\|\omega\|}. \quad (6)$$

Its mathematical expression is as follows:

$$\min_{\omega, b} \frac{1}{\|\omega\|}, s, t, y_i (\omega^T z_i + b) \geq 1, i = 1, 2, \dots, n. \quad (7)$$

The above problems can be equivalently transformed into the following problems:

$$\min_{\omega, b} \frac{1}{2} \|\omega\|^2, s, t, y_i (\omega^T z_i + b) \geq 1, i = 1, 2, \dots, n. \quad (8)$$

The constraints are integrated into the objective function through the Lagrangian function, as follows:

$$L(\omega, b, a) = \frac{1}{2} \|\omega\|^2 - \sum_{i=1}^n \alpha_i (y_i (\omega^T x_i + b) - 1). \quad (9)$$

Then make

$$\beta(\omega) = \max_{a_i \geq 0} L(\omega, b, a). \quad (10)$$

Our question is transformed into a request:

$$\min_{\omega, b} \beta(\omega) = \min_{\omega, b} \max_{a_i \geq 0} L(\omega, b, a). \quad (11)$$

Further transformed into

$$\max_{a_i \geq 0} \min_{\omega, b} L(\omega, b, a). \quad (12)$$

In order to solve equation (12), we first find the minimum values of  $\omega$  and  $b$  in equation (9), and calculate the extreme values of  $\omega$  and  $b$ , respectively, according to the extreme value conditions, namely:

$$\begin{aligned} \nabla_{\omega} L(\omega, b, a) &= 0 \\ \nabla_b L(\omega, b, a) &= 0. \end{aligned} \quad (13)$$

Solve to get

$$\omega = \sum_{i=1}^x \beta_i \alpha_i y_i. \quad (14)$$

$$\sum_{i=1}^n \alpha_i y_i = 0. \quad (15)$$

Substituting (14) and (15) into (12), the dual form of the original problem is obtained as follows:

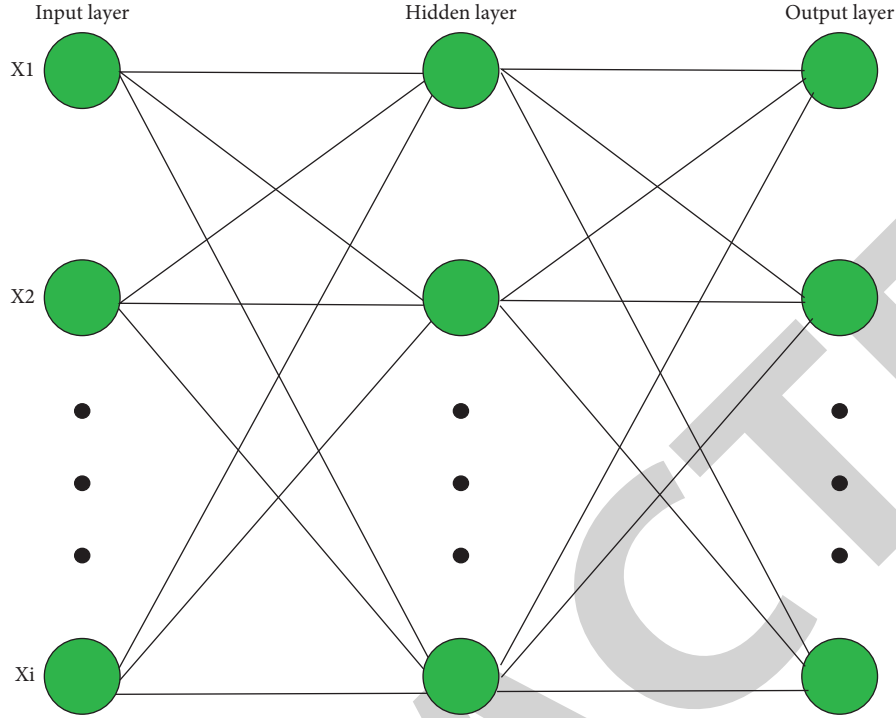


FIGURE 3: Multilayer feedforward neural network.

$$\max \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j \quad f\Phi^2(X)dx < \infty. \quad (16)$$

Then, there are

$$t, s, \alpha_i \geq 0, i = 1, 2, \dots, n, \sum_{i=1}^n \alpha_i y_i = 0. \quad \iint K(z, z')\Phi(z')dzdz' > 0. \quad (22)$$

That is, the original nonlinear data are mapped to a new feature space through a nonlinearity, and the following function is obtained:

$$f(Z) = \sum_{i=1}^n \omega_i \varphi_i(Z) + b. \quad (17)$$

Assuming that the data can be obtained by the linear combination of training points, it can be expressed as the following formula:

$$f(Z) = \sum_{i=1}^n \alpha_i y_i \langle \varphi(Z_i) \cdot \varphi(Z) \rangle + b. \quad (18)$$

Assuming that the kernel is a function  $K$ , for all  $x, z \in x$ , satisfy

$$K(X, Z) = \langle \varphi(x) \cdot \varphi(z) \rangle. \quad (19)$$

Then,  $\Phi$  is the mapping from  $X$  to the inner product feature space  $F$ . The Gaussian kernel  $K$  is

$$K(z_i, z_j) = \exp\left(-\frac{\|z_i - z_j\|^2}{2\sigma^2}\right). \quad (20)$$

That is, for any  $\Phi(x) \neq 0$ , there is

We turn the linear inseparable problem into

$$\max_a \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j K(x_i, x_j) \quad (23)$$

$$t, s, \alpha_i \geq 0, i = 1, 2, \dots, n, \sum_{i=1}^n \alpha_i y_i = 0.$$

By solving it, we can get an optimal SVM algorithm optimization method so as to improve it.

### 2.3. Distributed Collaborative High-Performance Computing.

The collaborative computing of the distributed model refers to the realization of simulation applications by coordinating and collaborating in a distributed environment with the help of computer, network technology, and model resources. This service sharing-based model not only enables the model to run in the best operating environment and ensures the safety of the model code, but also uses more computing resources and storage space to complete various complex application simulations. Using more computing resources and storage space to complete various complex application simulations can better protect the simulated data. Therefore, DCHF-SI needs to provide a collaborative environment for remote coupling and interaction of physical models located in different locations. The interaction process of the model is

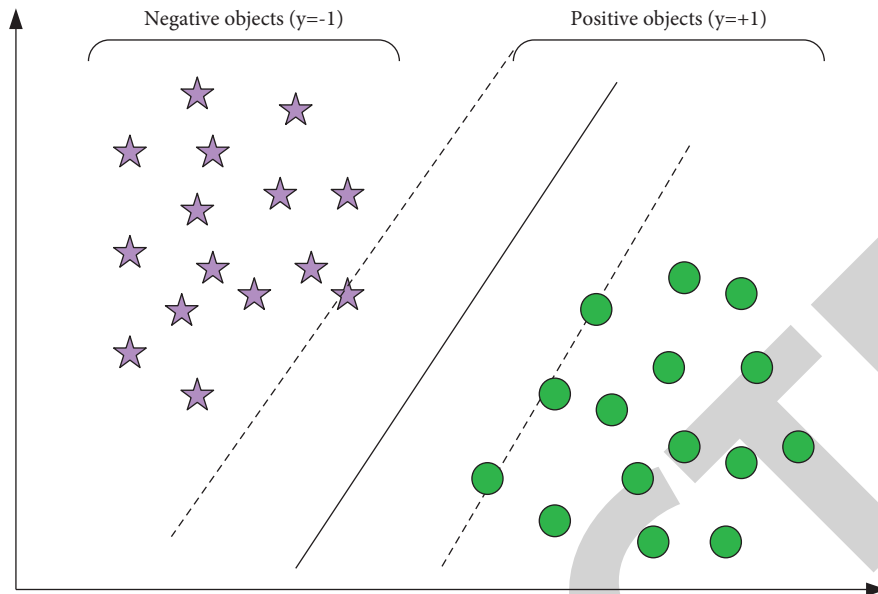


FIGURE 4: An example graph of 2-dimensional separability.

transparent to the user and needs to meet the particularity of different spatial physical applications and general commercial applications; the structure of DCHF-SI is shown in Figure 5.

**2.3.1. Coupling Interaction Basic Framework Layer.** The coupling interaction mechanism is responsible for the interaction coupling and communication of remote model components. The coupled interaction basic framework is the core supporting platform of the entire framework operation [20].

**2.3.2. Framework Service Component Layer.** Fault-tolerant service components provide checkpoints for setting up network or grid resources and rollback recovery and other fault tolerance strategies and log strategies [21].

**2.3.3. Physical Model Component Layer.** The physical model components are mainly obtained by encapsulating the physical model.

**2.3.4. Graphical User Interface (GUI) Layer.** Model sharing provides researchers with a large amount of model resources. However, it is very difficult for researchers to select the model they need from the massive model resources, and it is impossible to study and analyze all the available models. In order to solve this problem, the screening of related resource models is essential. As shown in Figure 6, when the model component is coupled with the internal model, there are  $M * N$  assembly methods, which makes the application assembly more complicated [22].

Therefore, this paper proposes a new type of component-abstract domain component. Models in the same domain share an abstract domain component. It also has a unified

standard port for this type of model, but it is not responsible for its implementation; instead, the best model components are selected through the QOS service to participate in the calculation, thereby achieving transparent and intelligent assembly of applications and reducing the burden on developers. The following interfaces need to be added to realize the generation and connection of abstract components [23]. The design scheme of the QOS service customizable northbound interface describes the external form and the internal function model of the interface, and based on this scheme, a northbound interface that provides bandwidth and priority customizable services is implemented. The interface is tested in the SDN experimental network, and the test results show that the scheme can achieve the expected goal.

### 3. Investigation and Experiment on the Status Quo of Urban Communities

A city is a product formed to adapt to the times, a manifestation of mature economic development, and a settlement formed by a multifaceted population. As the center of trade, it continues to attract foreigners, the number of permanent residents also increases, and coupled with resource constraints and unreasonable urban layout, it has been greatly affected. In order to improve this situation, government departments have established urban communities. The specific implementation steps are: first, delineate a range, and then build concentration camps in this area, such as commercial housing communities, office communities, etc., to guide relevant people to reside in them, to save resources as much as possible, use limited space to house more people, and let them form a society with a large number of people and a high concentration of people. Housing in the city community is mostly engaged with nonagricultural work, which can accommodate people from various places and



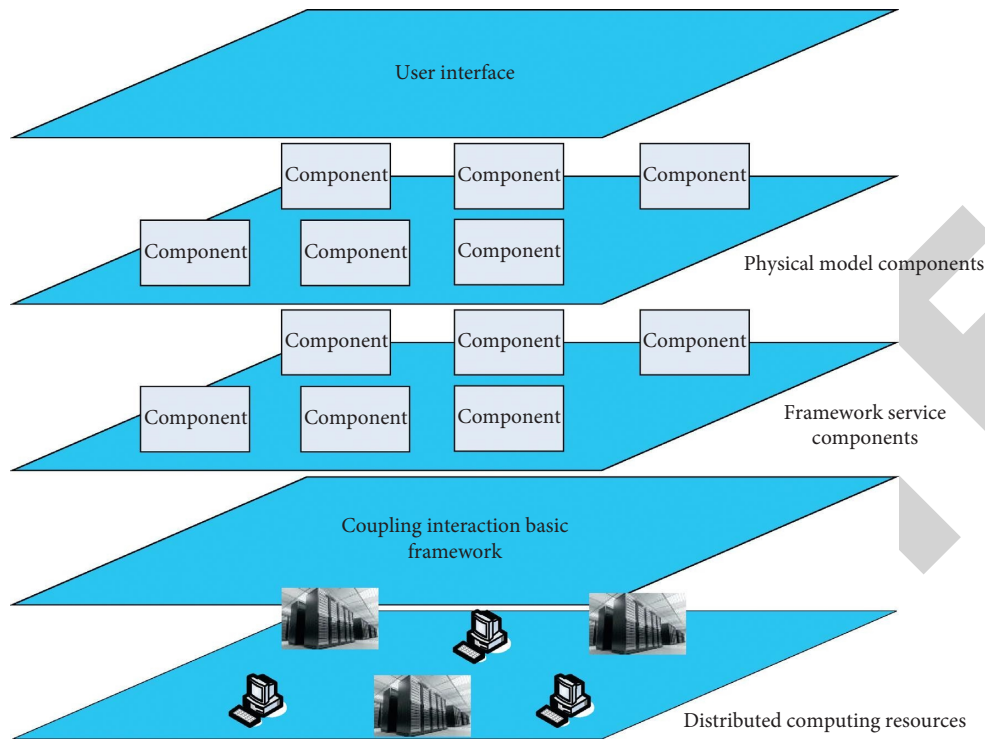


FIGURE 5: DCHF-SI structure diagram.

various types of work, and the rational use of space to create a small, densely populated social group. However, compared with rural areas, there is a lack of interpersonal communication [24].

At present, China's urban community governance has made great progress in terms of systems and mechanisms, but there are still many problems. There are many reasons for these problems, including the inertial effect of the original system, the influence of cultural traditions, and the lag in social development. For these issues cannot be ignored, it is necessary to explore the outstanding performance in different practices. This chapter is based on the status quo of community governance in urban communities, through combing the status of community governance practices, and discovering some potential problems in urban community governance so as to obtain beneficial enlightenment and help improve urban community governance. Here, we make a rough statistics of the number of communities within the scope of the survey statistics, as shown in Table 1:

It can be seen from the statistical data in the above table that social organizations have grown rapidly in the past two years, and the number of private units has grown relatively slowly; while foundations have no growth opportunities, the rapid growth of social organizations will lead to a management problem. If there is no good management and governance method, it is easy to cause problems. Therefore, a survey of satisfaction with the governance methods of urban community groups is added, as shown in Table 2:

Obviously, it can be seen that the problem caused by the excessively rapid growth of social organizations is that there are many governance problems, governance chaos, and the

satisfaction of group members is low, while private units basically grow slowly, showing satisfaction and dissatisfaction, and the satisfactory conditions are basically the same. The foundation has grown by one in three years, and this has resulted in them having enough time to administer the foundation, so that the vast majority of people can be satisfied, and the satisfaction will naturally increase [25].

**3.1. Weak Community Social Organizations.** In the context of modern development, urban community governance, such as the introduction of the participation of social forces, can add new vitality to community services. Social organizations participating in community governance can give full play to their role as a link, integrate their own resource advantages, broaden funding sources for community governance, assume part of the governance capacity, and use their professional capabilities to provide residents with market-oriented services.

A social survey experiment on the degree of understanding of community social organizations was organized here, and the experimental data obtained are shown in Table 3:

From the above table, we can see that in urban communities, children's understanding of community social organizations accounted for 2%, adults accounted for 8%, and the elderly 2%. It only accounts for 12% of the total, and the number of people who understand it is really too small. Children and the elderly have less knowledge of community social organizations, and it is mainly concentrated that adults have more knowledge of urban community organizations. As a result, most people have problems with the

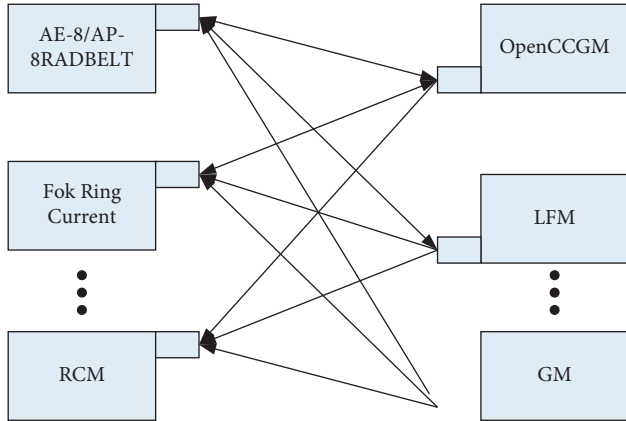


FIGURE 6: Model component and internal model component coupling description.

TABLE 1: Statistics on the number of urban community groups.

Group	Social groups	Private unit	Foundation
2016	75	32	9
2017	91	38	9
2018	114	43	10

TABLE 2: Satisfaction of urban community groups.

Group	Social groups	Private unit	Foundation
Satisfy	56	221	389
Dissatisfied	489	243	31

governance of the community. This also reflects the problems of small-scale social organizations, insufficient development, and insufficient influence in Chinese cities. Community funding is not guaranteed, and it is easy to be administrated. The inability to independently assume the responsibility of community management has led to inadequate governance in the community [26].

**3.2. Residents Lack Awareness of Active Participation.** Serving community residents is the purpose and goal of community governance, and all governance actions are based on safeguarding the interests of the residents. The extent of residents' participation in community autonomy to a large extent reflects the true level of community autonomy. From the survey of urban community residents participating in community activities in Table 4, it can be seen that the basis for the formation of this sense of identity and belonging is gradually disintegrating [27].

Analyzing the data in Table 4, we can see that the proportion of children who often participate is 4%, adults 6%, and the elderly 2%. These values are far lower than those of a healthy and qualified urban community. The proportion of participants is only 12%, the proportion of those who have participated in some is 30%, the proportion of those who have participated occasionally is 33%, and the proportion of those who have never participated is 25%. Similarly, some seldom participate in community activities, indicating that

there are fewer feedbacks on problems in this community, and people cannot conduct a good and effective management of the urban community. For this reason, we have also made a detailed survey on the reasons why people of different ages do not participate in community activities. The survey data are shown in Table 5:

From Table 5, we can still see that adults still pay attention to the interests of the majority, and this is one of the reasons why most people do not participate in community activities. The reason for this phenomenon is that the community's publicity is still insufficient, thus failing to make people develop the concept that community governance is related to everyone. This has also caused more and more people to lose confidence in community management.

Another questionnaire is designed here to investigate and reform the community governance system. There are 21 questions in the questionnaire, which consist of two parts: the first part is the basic information of the interviewee, covering nine aspects including gender, age, marital status, location of household registration, education level, residence time, identity, political outlook, and average monthly income. It forms a characterization description of the interviewee; the second part is a survey of the current community governance system. Questions include community party organizations, community neighborhood committees, community service stations, community intermediary organizations, residents' participation, community-related policies, and other elements of the community governance system.

## 4. Status Quo and Improvement of the Community

**4.1. Relevant Situation of the Community Governance System.** From the selection of residents, it can be found that almost ordinary people regard the community committee as the first choice to help them solve the difficulties in daily life, which shows that the community committee has a relatively high sense of identity in the hearts of residents, followed by government departments and community service stations. Compared with the other three governance bodies, the owners' committee and the property company do not seem to be "rooted" (the other three governance entities mainly include: private units, foundations, and social groups), which also reflects the absence of the two in community governance, which is not conducive to building a "social coordination" community governance system in the long run. The relevant situation investigation is shown in Figure 7:

From the table above, we can see that in urban communities, the number of residents who often live with neighborhood committees accounts for about 15%. Among them, children and the elderly account for more, while adults still account for a minority. About 40% of the people who don't care about the governance and management of the community and occasionally contact the neighborhood committee account for about 40%. Among them, children and adults account for a relatively large number, while the elderly are relatively small. Among those who almost never

TABLE 3: Urban community residents' understanding of community social organizations.

Generation	Very familiar (%)	Probably understand a little (%)	Don't understand at all (%)
Child	2	8	18
Adult	8	31	2
The elderly	2	10	17

TABLE 4: Urban community residents' participation in community activities.

Condition	Child (%)	Adult (%)	The elderly (%)
Often attend	4	6	2
Participated in some	10	17	3
Participate occasionally	8	20	5
Never participated	7	15	3

TABLE 5: Participation in community activities by people of different ages.

Condition	Child (%)	Adult (%)	The elderly (%)
No benefit	1	30	4
No time to	2	12	3
No effect	3	15	5
Don't know to Participate	8	1	3
Never heard of that	6	3	4

contact the neighborhood committee, adults account for the overwhelming majority, while children and the elderly account for only a small proportion. In the statistics of familiarity, the proportion of people who know the neighborhood committee is about 25%. On the contrary, the proportion of children is the largest, while the proportion of adults and the elderly is relatively small. Among those who do not know the director of the neighborhood committee at all, children and adults account for a larger proportion, while the elderly are fewer. Among those who have heard of it but are not clear, there are more adults and fewer elderly and children. The statistics on the familiarity of community party organizations and the degree of recognition of their related governance methods are shown in Figure 8:

From the survey statistics in Figure 9 on the familiarity of community party organizations and the recognition of related governance methods, we can see that the managerial personnel in the community neighborhood committees still have a relatively high percentage of recognition of the governance methods of urban communities, reaching 15.9%; in addition, the managers of all organizations have a high degree of recognition of the governance methods, and the members of them generally do not agree with the governance methods, and they are maintained at a low level. The senior citizens who agree with party secretaries are the elderly, and most adults don't know the secretaries of the community's party organizations. When the main personnel in an urban community do not know the secretaries of the community's party organization, it can only indicate the urban community. The publicity of the community is still not in place, resulting in an insufficient sense of participation of community members and the inability to actively participate in the comprehensive governance and control of the community.

In summary, the "two committees" of the community are not closely connected to the residents, and the community party organizations and community neighborhood committees are precisely the "top people" who drive residents to actively participate in community governance. Therefore, the construction of the "two committees" in the community governance system needs to be further improved and strengthened. At the same time, the problem activities organized by the community are counted, as shown in Figure 9:

From the urban community's participation in community activities organized by the community, we can see that most of the adults in cultural and sports activities have not participated in any of the activities. The main reason is to let their children participate in some activities organized by the community, but if it's just like this, it means that they have not participated in community team activities at all. Children's participation in community activities does not have much effect on the progress of the community. Similarly, in terms of public welfare activities, the majority of children participate in activities, followed by the elderly who stay at home, and most adults do not participate in the governance of the community, which reflects the fact that the activities of the community mainly remain at the level of formality, and there is no in-depth understanding of the meaning of community activities.

From the above data results, we can find that most residents are not very active in participating in community activities, and community activities are one of the effective ways to increase the sense of belonging and identity of community residents to the community. Therefore, how to improve the community governance system and eliminate the gap in public participation is worthy of our deep consideration.

#### 4.2. Distributed High-Performance Computing Capability.

Analyze the phenomenon in the process of urban governance through the calculation and analysis of the surveyed data, and analyze the problems of urban community governance. Therefore, it is possible to deduce a most suitable governance method, cooperate with community managers to better manage the community, and creatively explore the most suitable governance methods for each urban community.

Here, we combine the experimental data investigated by the above analysis and analyze the community governance issues through distributed high-performance computing, coordinate the role of various organizations in urban communities, and actively guide community members to participate in the governance of the community. The analysis data are shown in Figure 10:

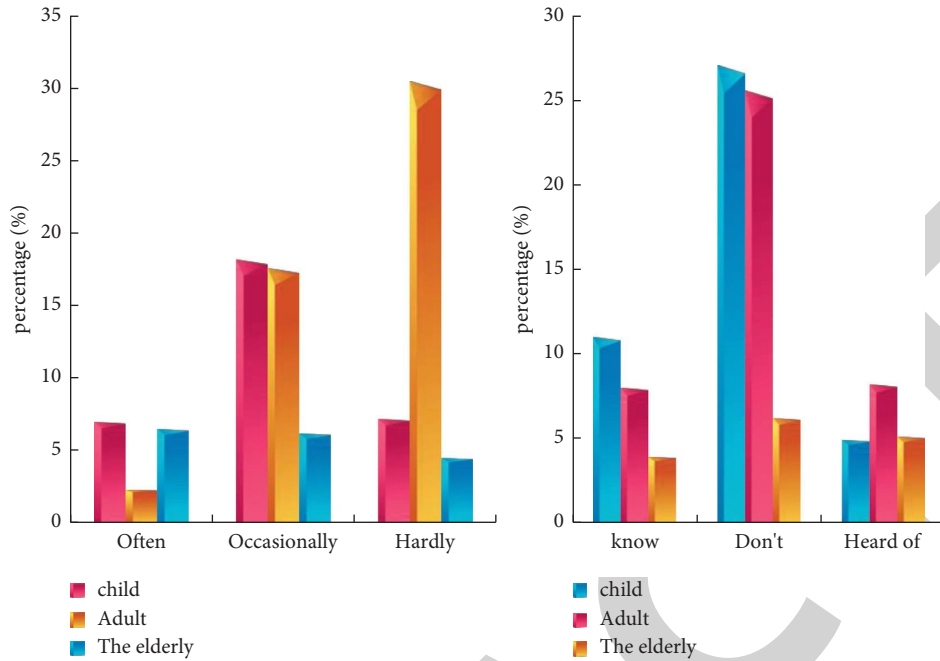


FIGURE 7: Statistics on the frequency and familiarity of residents' contact with neighborhood committees.

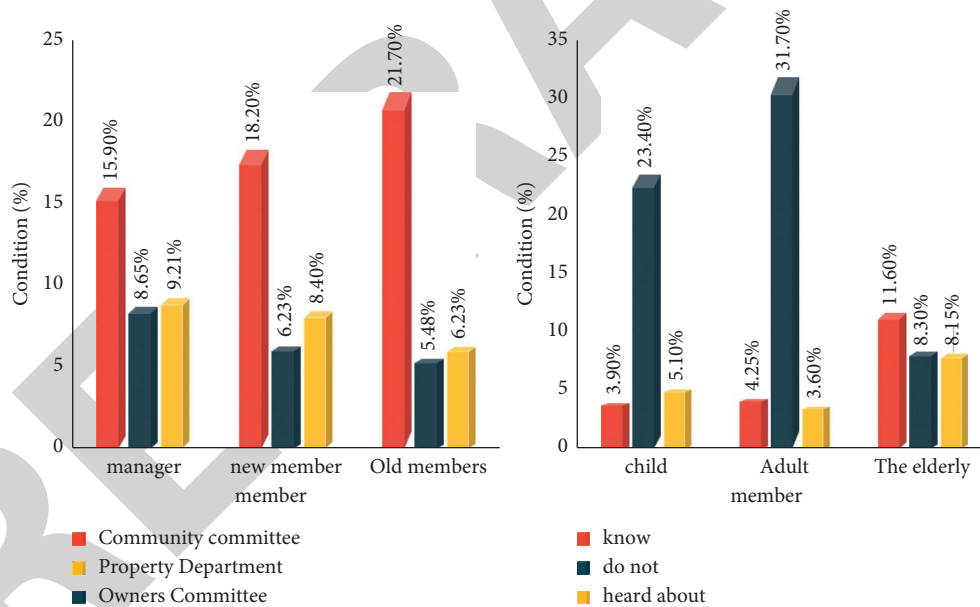


FIGURE 8: Statistics on the familiarity of community party organizations and the degree of recognition of their related governance methods.

From the above Figure 10, we can see that the development distribution curve of the community is more concentrated, which is more conducive to the stable construction of the urban community. Regarding the exponential growth of the community, the environment of the urban community has increased by 21.1% compared with previous years, and the community is safer than in previous years. With an increase of 11.1%, the development of urban communities has increased by 26% compared to previous years. The overall situation shows growth; it can be seen that using distributed high-performance computing to calculate

the various factors of the community, the most suitable community development plan can be deduced, which is more conducive to the development of the community. After implementation, the overall urban community governance has increased by 19.4%.

The data obtained through calculation and analysis is more conducive to our assumption of a sound and harmonious urban community and the establishment of a democratic rule of law, fairness and justice, honesty and friendship, full of activities, stability and order, and harmonious coexistence between people's living environments.

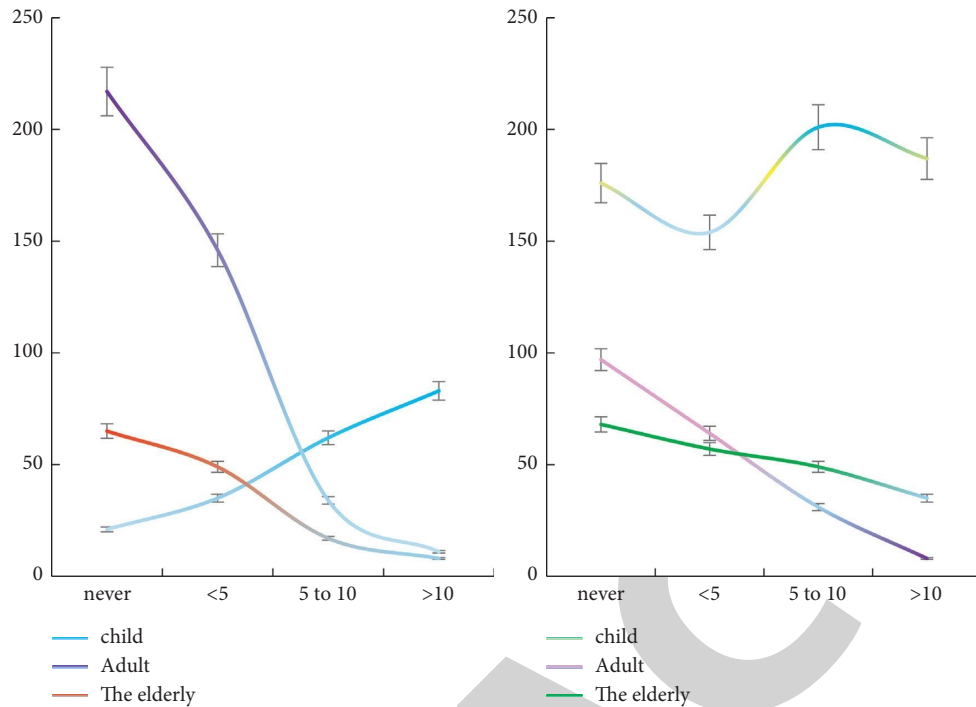


FIGURE 9: Residents' participation in cultural and sports activities and public welfare activities organized by the community.

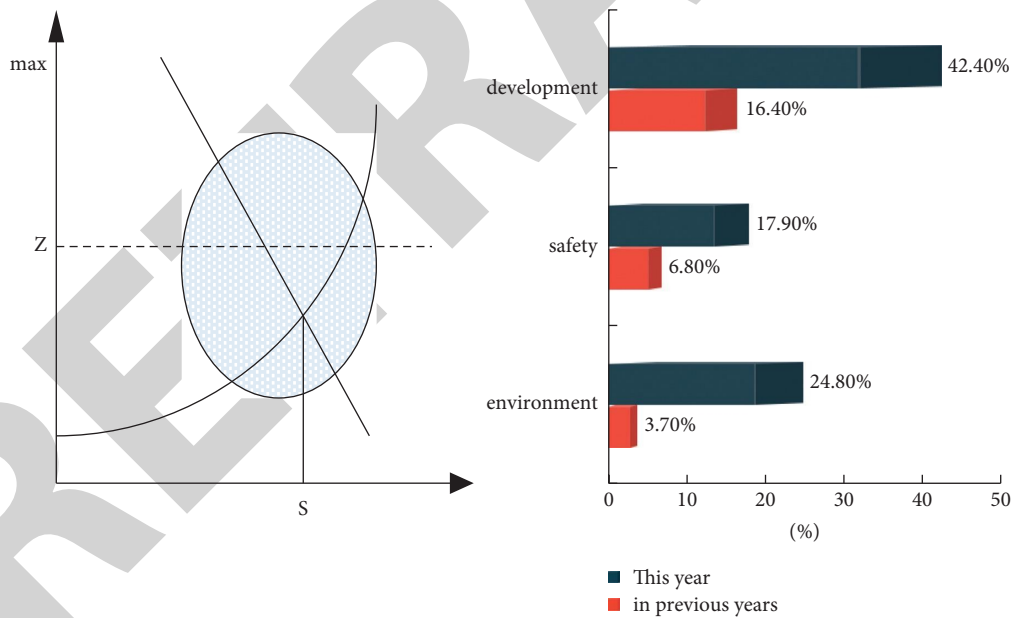


FIGURE 10: The distribution curve of urban community development and the exponential growth of all aspects of the community.

### 5. Conclusions

This paper mainly conducts innovative research on the current domestic urban community governance methods, mainly using the computing power of distributed algorithms and the use of key algorithms of blockchain technology, and comprehensively analyzes and understands the current community through the combined use of the two. The goal of improving the current domestic community management

dilemma has been achieved by exploring the management methods and improving the management methods. At the same time, the analysis and comparison of the urban community governance mechanism after the analysis of distributed high-performance computing is carried out, and it is found that through the calculation derivation the proposed governance method can improve the overall governance effect of the community by 19.4%. It can be seen that the use of related distributed algorithms in this paper is

## Research Article

# A Boosting-Based Deep Distance Metric Learning Method

Zilong Li <sup>1,2,3</sup>

<sup>1</sup>*School of Information Engineering, Xuzhou University of Technology, Xuzhou 221018, China*

<sup>2</sup>*School of Computer Science & Technology, China University of Mining and Technology, Xuzhou 221116, China*

<sup>3</sup>*Post-Doctoral Research Center, Onnes Power Technology Co., Ltd., Xuzhou 221003, China*

Correspondence should be addressed to Zilong Li; [lzl@xzit.edu.cn](mailto:lzl@xzit.edu.cn)

Received 30 January 2022; Revised 21 February 2022; Accepted 25 February 2022; Published 14 March 2022

Academic Editor: Daqing Gong

Copyright © 2022 Zilong Li. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

By leveraging neural networks, deep distance metric learning has yielded impressive results in computer vision applications. However, the existing approaches mostly focus a single deep distance metric based on pairs or triplets of samples. It is difficult for them to handle heterogeneous data and avoid overfitting. This study proposes a boosting-based learning method of multiple deep distance metrics, which generates the final distance metric through iterative training of multiple weak distance metrics. Firstly, the distance of sample pairs was mapped by a convolution neural network (CNN) and evaluated by a piecewise linear function. Secondly, the evaluation function was added as a weak learner to the boosting algorithm to generate a strong learner. Each weak learner targets the difficult samples different from the samples of previous learners. Next, an alternating optimization method was employed to train the network and loss function. Finally, the effectiveness of our method was demonstrated in contrast to state of the arts on retrieving the images from the CUB-200-2011, Cars-196, and Stanford Online Products (SOP) datasets.

## 1. Introduction

In the past decades, distance metric learning has been applied effectively in image retrieval, face recognition, person re-identification, clustering, etc. It is now a hot topic in the field of computer vision. Thanks to the recent success of convolutional neural networks (CNNs), deep distance metric learning methods have attracted lots of attention [1].

Each deep distance metric aims to map training samples to features via CNNs. The mapping should narrow the distance between similar sample pairs and increase that between dissimilar sample pairs. To learn deep distance metrics, many approaches have been developed based on sample pairs [2, 3], triplets [2, 4], or quadruplets [5]. This study attempts to learn the simple similarity functions of sample pairs. The distance metric was defined as the Euclidean distance between sample pairs, which can be computed rapidly compared with other metrics.

Most of the existing methods of deep distance metric learning try to improve the single loss function based on a single distance metric. However, a single distance metric is insufficient to handle all the samples from the given data

distribution. In fact, feature data are generally not distributed uniformly: the density varies from region to region in the data distribution [6, 7]. To solve the problem, some scholars resorted to ensemble technique and employed several learners to map each sample to multiple subspaces [8–10]. Nevertheless, these strategies do not support end-to-end training of the network and loss function of each weak learner. The lack of this training model suppresses the discrimination ability and increases the susceptibility of the metric to noise. The accuracy of deep distance metric learning could be further improved through joint training of the network and loss function.

This study aims to improve the adaptability of conventional deep distance metric between pairs of samples. The main idea is to divide the last fully connected layer of the CNN into multiple nonoverlapping groups (Figure 1), each of which is a separate feature mapping of the network. The distance metric of sample pairs mapped by one of the groups was evaluated by a piecewise linear function. Each group has a corresponding evaluation function, which is added as a weak learner to the boosting algorithm to generate a strong learner. This finally forms a multidistance metric ensemble.



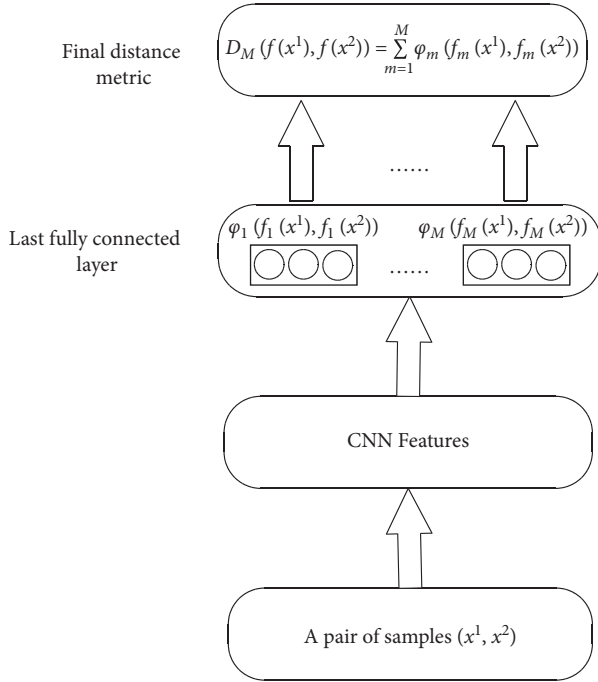


FIGURE 1: Structure of boosting-based deep distance metric.

In addition, the same underlying feature representation, which was pretrained through experiments, was applied to the fully connected layers of all groups. In this way, the high computing cost of CNN training in the boosting framework was significantly reduced. In that framework, each learner reweights the training samples for successive learners, according to the gradient of the loss function. As a result, the successive learners would focus on difficult sample features, producing more suitable feature representations. The final ensemble output is a linear composition of multiple weak learners. Furthermore, the performance of the conventional distance metric was improved by introducing a piecewise linear function, which evaluates the similarity of sample pairs in distance metric learning. This facilitates the joint training of the network and loss function. Through the evaluation of various deep distance metric learning methods in the image retrieval task, it can be seen that Recall@1 of the proposed method is 4.2, 2.8, and 0.4 higher than that of the previous best score on CUB-200-2011, Cars-196, and SOP datasets, respectively. Experimental results show that the proposed method outperforms the comparison methods, while avoiding overfitting to a certain extent.

The contributions of our work are as follows:

- (1) The last fully connected layer of the CNN was used to form multiple groups of features, which was designed to form a distance metrics ensemble and formulated as a boosting problem. Then, an alternating optimization method was adopted to jointly train the network and loss function.
- (2) A piecewise linear function was employed as the evaluation function of the distance metric of sample pairs mapped by CNN and added as a weak

learner to the boosting algorithm to generate a strong learner.

## 2. Literature Review

This section reviews the most closely related works out of the numerous publications on the hot topic of distance metric learning.

**2.1. Deep Distance Metric Learning.** Many methods employ a discriminative distance metric loss function to increase interclass distance and reduce intraclass distance [2, 4, 11–14]. For instance, contrastive loss is a popular tool of deep distance metric learning that minimizes the distance between the eigenvectors of positive sample pairs and widens the distance between the negative sample pairs [2, 3]. Based on contrastive loss, triplet loss creates a 3-tuple with a positive sample pair and a negative sample pair, in the light of the relative relationship between intraclass distance and interclass distance [2, 4], and ensures that the positive sample pairs are closer in the mapped feature space than the negative sample pairs. Many other loss functions have been extended from the above two losses, namely, histogram loss [15], quadruplet loss [5], N-pair loss [11], angular loss [12], and hierarchical triplet loss [16].

Taking a tuple of samples as training samples yields a huge amount of training data. Deep distance metric learning would be greatly enhanced by acquiring more effective samples. Recently, several scholars have designed sampling strategies to tackle hard and semihard negative mining [16–18]. For example, Xuan et al. [7] observed that easy positive samples help to preserve the intraclass difference and thus improve the generalization ability of triplet loss. However, the use of easy positive samples constantly underchallenges the metric, making the embedding space less discriminative.

**2.2. Ensemble Learning.** The methods above all strive to improve the loss function based on a single distance metric. However, it is difficult for them to adapt to all available data. Recently, ensemble learning, which iteratively trains an ensemble from several weak learners for the final prediction, has been incorporated to boost the generalization performance of deep metric learning.

Negrel et al. [8] explained how to use their boosting-based metric learning algorithm to compute hierarchical organizations of face databases. Kim et al. [9] introduced multiple attention-based learners for ensemble. Xuan et al. [10] grouped labels randomly to create a large family of related embedding models, which can serve as an ensemble. Sanakoyeu et al. [6] employed a divide-and-conquer strategy to divide the embedding space to several clusters and used each cluster to train a single learner.

**2.3. Other Related Metric Learning.** In addition, there are other types of distance metric approaches recently, such as sample selection, local metric, and hierarchical metric. Wu

et al. [19] proposed a distance-weighted sampling procedure, which selects more informative and stable examples than traditional approaches, achieving excellent results in the process. Wang et al. [13] generalized tuple-based losses and reformulated them as different weighting strategies of positive and negative pairs within a minibatch. Roth et al. [20] proposed to learn the distribution for sampling negative examples instead of using a predefined one. Local metric learning methods [21, 22] learn a collection of Mahalanobis distance metrics, each operating on a different subset of the data obtained by K-means or Gaussian mixture clustering. From [23, 24], we learn a two-level category hierarchy by using coarse and fine classifiers. Ge et al. [16] proposed a hierarchical version of triplet loss that learns the sampling all together with the feature embedding.

Different from the above approaches, our approach realizes the end-to-end training of the network and loss function of each weak learner, thereby enhancing the accuracy of deep distance metric and reducing the probability of overfitting.

### 3. Methodology

**3.1. Boosting-Based Deep Distance Metric Model.** Let  $X = [x_1, x_2, \dots, x_N]$  be  $N$  training sample pairs, each of which belongs to one of the two class labels  $y_n \in \{-1, +1\}$ . If the two samples belong to the same class, the pair is labeled  $y_n = +1$  and called a positive sample pair; if the two samples belong to different classes, the pair is labeled  $y_n = -1$  and called a negative sample pair.

We divide the last fully connected layer of the CNN into multiple nonoverlapping groups. The training sample pair  $x_n = (x_n^1, x_n^2)$  is fed into the CNN to generate an eigenvector pair  $f_m(x_n) = (f_m(x_n^1), f_m(x_n^2))$ , which is extracted from the  $m$ th group of the last fully connected layer. So, the training sample pair  $x_n$  can be mapped to generate multiple groups of features, which was designed to form a distance metrics ensemble and formulated as a boosting problem.

Drawing on the idea of the boosting algorithm, multiple weak learners are adopted to produce a strong learner of distance metrics between the mapping values of training sample pairs. The weak learners are trained on reweighted samples, according to the gradient of the loss function. In general, we want to a set of weak learners and their corresponding boosting model:

$$D_M(f(x_n^1), f(x_n^2)) = \sum_{m=1}^M \varphi_m(f_m(x_n^1), f_m(x_n^2)), \quad (1)$$

where  $M$  is the number of weak learners and  $\varphi_m$  is the distance metric evaluation function between the eigenvectors of the training sample pair mapped by the  $m$ th group of the fully connected layer.

In the above formula,  $\varphi_m$  was used to quantify the similarity between two training samples, and it responds to this similarity based on whether the two samples should be considered to represent the same class. Therefore, a

threshold was defined to deal with the distance metric between two training samples, and a piecewise linear function was adopted as the evaluation function. This function reduces the distance between similar training samples and increases that between dissimilar ones in the mapped space. The evaluation function can be defined as

$$\varphi_m(f_m(x_n^1), f_m(x_n^2)) = \begin{cases} \alpha_m, & \text{if } d(f_m(x_n^1), f_m(x_n^2)) < t_m, \\ \beta_m, & \text{else,} \end{cases} \quad (2)$$

where  $d(f_m(x_n^1), f_m(x_n^2))$  is a generic distance metric (the simple Euclidean distance),  $\alpha_m$  and  $\beta_m$  are the evaluated similarity and dissimilarity between the two samples, respectively, and  $t_m$  is a distance metric threshold. If the Euclidean distance between two mapped training samples is smaller than the threshold  $t_m$ , then the evaluation value is  $\alpha_m$ ; otherwise, it is  $\beta_m$  (Figure 2).

In each round of boosting, a new weak learner is trained on the reweighted training set in the minibatch, according to the gradient of the loss function, and then added to form a strong learner. As demonstrated by Friedman [25], the training of a single learner can be written as a loss function minimization problem:

$$\arg \min \sum_{i=1}^N \ell(y_i; D_{M-1}(x_i) + \varphi_m(x_i)), \quad (3)$$

where  $\ell$  is a loss function. Here, the exponential loss function  $\ell(y_i; D_M(x_i)) = e^{-y_i D_M(x_i)}$  is utilized. Inspired by Schapire and Singer [26], formula (3) can be rewritten as

$$\arg \min \sum_{i=1}^N w_i^m e^{-y_i \varphi_m(x_i)}, \quad (4)$$

where  $w_i^m$  is the weight of training sample  $x_i$  in iteration  $m$ . The weak learner is selected to minimize the loss function in each iteration to update the strong learner. Both  $\alpha_m$ ,  $\beta_m$ , and  $t_m$  of the distance evaluation function and  $w_m^{\text{net}}$  of the  $m$ th group of the fully connected layer need to be optimized.

The proposed approach is easily integrated into some deep metric learning approaches, such as triplet loss, N-pair loss, and hierarchical triplet loss. However, for some loss functions, such as histogram loss and angular loss, it is not applicable and needs to be improved.

**3.2. Joint Training.** As to formula (3), we need attempt to jointly learn both the network and loss function. We note that its function was nonconvex, which was difficult to solve in general. Referring to Zhang et al. [27], this study applies an alternating optimization method to jointly train the network and loss function.

Since a learner needs to be optimized in each round of boosting, the optimization problem (4) was investigated, while fixing parameters  $w_m^{\text{net}}$  of the  $m$ th group of the fully connected layer. Formula (4) can be decomposed into

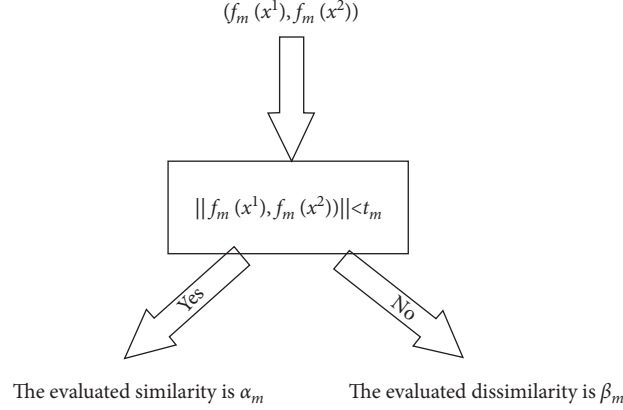


FIGURE 2: Operation of the evaluation function.

$$\sum_{i=1}^N w_i^m e^{-y_i \varphi_m(x_i)} = \sum_{y_i=1 \wedge d(x_i) < t_m} w_i^m e^{-\alpha_m} + \sum_{y_i=1 \wedge d(x_i) \geq t_m} w_i^m e^{-\beta_m} + \sum_{y_i=-1 \wedge d(x_i) < t_m} w_i^m e^{\alpha_m} + \sum_{y_i=-1 \wedge d(x_i) \geq t_m} w_i^m e^{\beta_m}. \quad (5)$$

Taking partial derivatives of formula (5) with respect to  $\alpha_m$  and  $\beta_m$  and setting both to zero to optimize each parameter,

$$\alpha_m = \frac{1}{2} \log \left( \frac{\sum_{y_i=1 \wedge d(x_i) < t_m} w_i^m}{\sum_{y_i=-1 \wedge d(x_i) < t_m} w_i^m} \right), \quad (6)$$

$$\beta_m = \frac{1}{2} \log \left( \frac{\sum_{y_i=1 \wedge d(x_i) \geq t_m} w_i^m}{\sum_{y_i=-1 \wedge d(x_i) \geq t_m} w_i^m} \right). \quad (7)$$

After each iteration, the weights of the training sample pairs are updated using the exponential loss function:

$$w_i^{m+1} = w_i^m e^{-y_i \varphi_m(x_i)}. \quad (8)$$

Then, the weights of all training sample pairs are normalized. As shown in formulas (6) and (7), the parameters that affect the evaluation function are only related to  $t_m$ , i.e., the optimal value obtained by the traversal method. If the training sample pairs are classified correctly, the weight of successive learners tends to be small; otherwise, the weight tends to be large. Hence, successive learners focus on different training sample pairs than previous learners, increasing the diversity among learners (Figure 3).

The next step is to update parameters  $w_m^{net}$  of the  $m$ th group of the fully connected layer, while fixing  $\alpha_m$ ,  $\beta_m$ , and  $t_m$  of the evaluation function. These parameters were trained with the contrastive loss function, using the standard backpropagation algorithm. In the forward process, the similarity distance metric is computed for each input training sample pair. In the backward process, the gradient of the loss function is iteratively propagated for each group (Figure 4).

For the contrastive loss function, the distance metric threshold  $t_m$  obtained through weak learning training serves as the distance margin of a negative training sample pair. Then, the contrastive loss function can be established as

$$\begin{aligned} \ell'(x_n^1, x_n^2) &= \frac{1}{2} (y_n + 1) d(f_m(x_n^1), f_m(x_n^2)) \\ &+ \frac{1}{2} (1 - y_n) [t_m - d(f_m(x_n^1), f_m(x_n^2))]_+. \end{aligned} \quad (9)$$

The training procedure is illustrated as Algorithm 1.

#### 4. Experiments and Results' Analysis

To verify its effectiveness, our method for deep distance metric learning was tested on three standard datasets: CUB-200-2011, Cars-196, and Stanford online products (SOP). Following the standard protocol proposed by Oh Song et al. [2], each dataset was broken down into a training set and a test set. For the CUB-200-2011 dataset, 5,864 images in the first 100 classes were allocated to the training set and 5,924 images in the last 100 classes were allocated to the test set. For the Cars-196 dataset, 8,054 images in the first 98 classes were allocated to the training set and 8,131 images in the remaining 98 classes were allocated to the test set. For the SOP dataset, 59,551 images of 11,318 classes were allocated to the training set and 60,502 images in 11,316 classes were allocated to the test set.

The performance of our method in retrieving images from the above datasets was evaluated by Recall@K. For each retrieval task, the authors computed the percentage of the testing images whose top-K retrieved images contain at least one image with the same class label. The K value was set to  $K \in \{1, 2, 4, 8, 16, 32\}$  for CUB-200-2011 and Cars-196, and  $K \in \{1, 10, 100, 1000\}$  for SOP. Our method was implemented under the framework of TensorFlow. Following Oh Song's approach [2], GoogLeNet was adopted as the feature extractor. The batch size was fixed at 128 in all experiments.

Since the deep distance metric could be affected by the number of weak learners, the influence of that number on our method was observed through experiments on each of the three datasets. As shown in Figure 5, with the growing number of weak learners, the Recall@1 score first increased

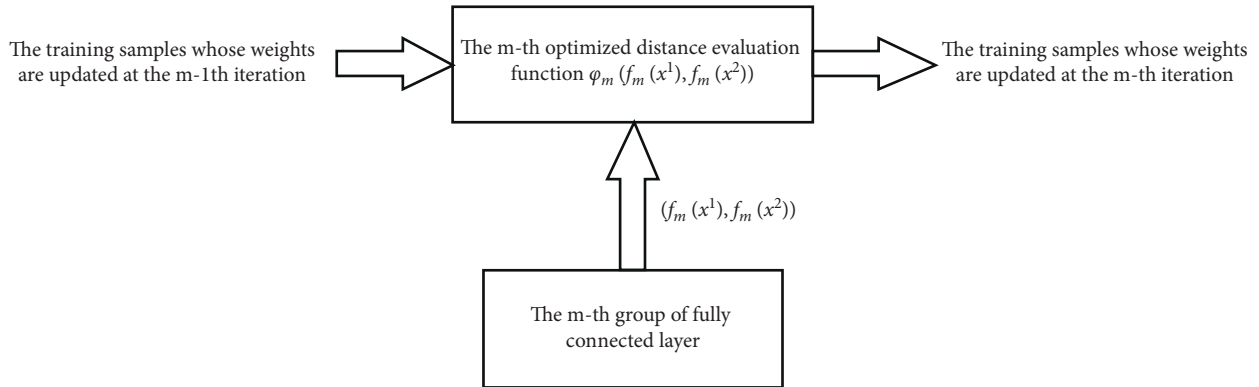


FIGURE 3: Optimization of the distance evaluation function with fixed parameters of the  $m$ th group of the fully connected layer.

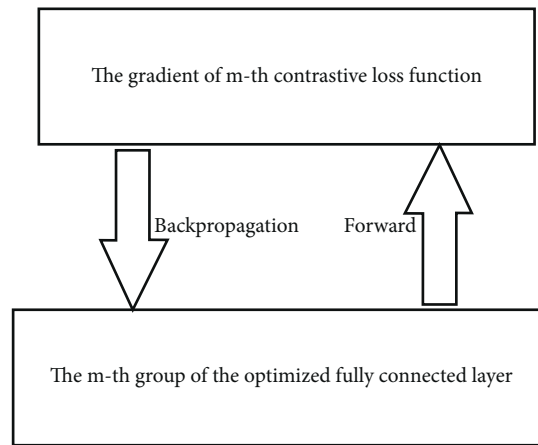


FIGURE 4: Optimization of the  $m$ th group of the fully connected layer with fixed parameters of the evaluation function.

```

Initialization: initialize distance metric threshold  $t_m$ ,  $m = 1, \dots, M$ , the weights of the training sample pairs  $w_n^m$ ,  $m = 1, \dots, M$  and  $n = 1, \dots, N$ ,
and parameters of the  $m$ th group of the fully connected layer  $w_m^{\text{net}}$ ,  $m = 1, \dots, M$ , randomly;
for  $m = 1$  to  $M$ 
  repeat
  for  $n = 1$  to  $N$ 
  Select the optimal value  $t_m$ , which minimizes formula (4), and update the weights  $w_n^m$  of the  $n$ th training sample pairs;
  Compute the derivatives of formula (9), and update the parameters  $w_m^{\text{net}}$  of the  $m$ th group of the fully connected layer via
  backpropagation;
  end
  until terminal condition.
 $D_m = D_{m-1} + \varphi_m$ 
end
Output  $D_M$ 

```

ALGORITHM 1: Training procedure of our method.

and then declined. The highest Recall@1 score was achieved at 8, 6, and 7 weak learners, for CUB-200-2011, Cars-196, and SOP, respectively. A possible reason is that the images in Cars-196 have relatively small variations, those in SOP face large view-point changes, and those in CUB-200-2011 feature a large pose variation and a strong background clutter. In the following experiments, the number of weak learners

was set to 8, 6, and 7, for CUB-200-2011, Cars-196 and SOP, respectively.

The eigenvector size also exerts a major effect on the deep distance metric. Hence, experiments were carried out on Cars-196 with 6 weak learners and different eigenvector sizes. Drawing on the work of Wang et al. [13], the eigenvector size was increased from 64 to 1,024. Figure 6

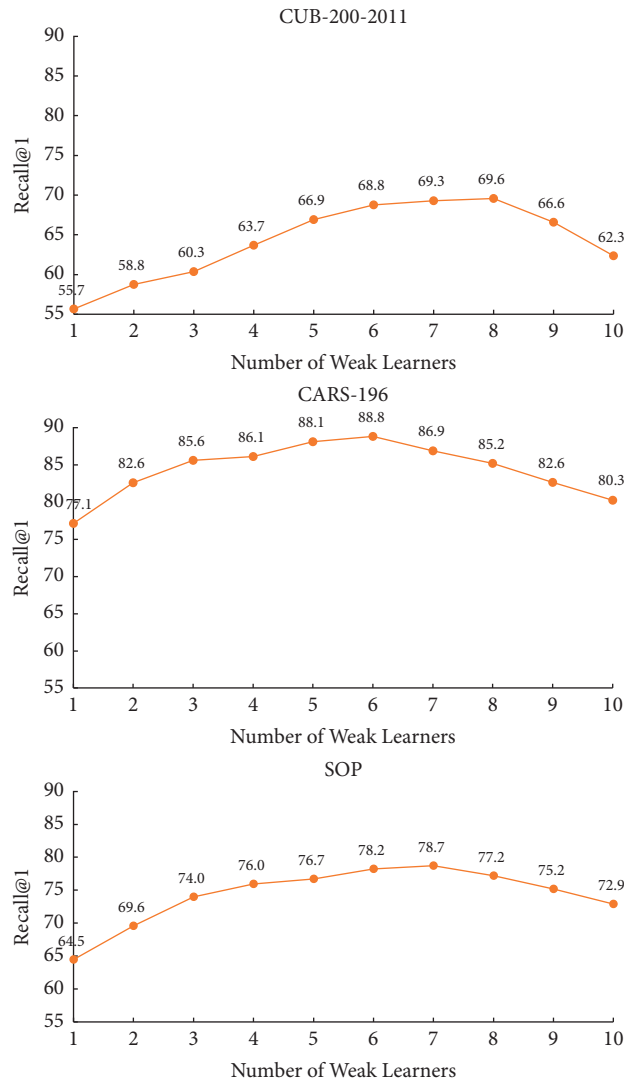


FIGURE 5: Influence of the number of weak learners over the performance of our method on different datasets.

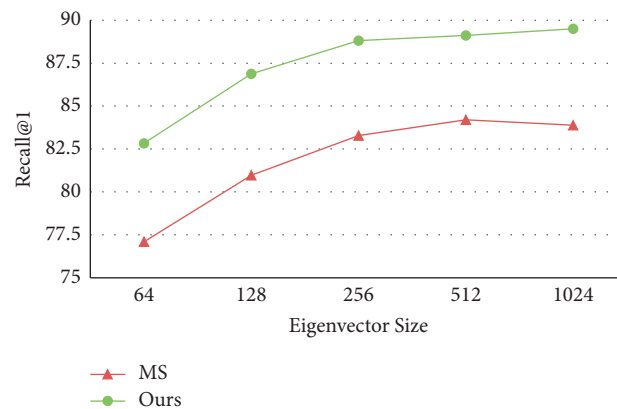


FIGURE 6: Retrieval performance (R@1) at different eigenvector sizes on Cars-196.

compares the retrieval performance of our method with that of the multiscale (MS) method [13]. As shown in Figure 6, the retrieval performance of both methods gradually increased with the eigenvector size. Our

method performed stably, when the size was equal to or greater than 256, and always outshined the MS. Hence, the eigenvector size was fixed at 256 in subsequent experiments.

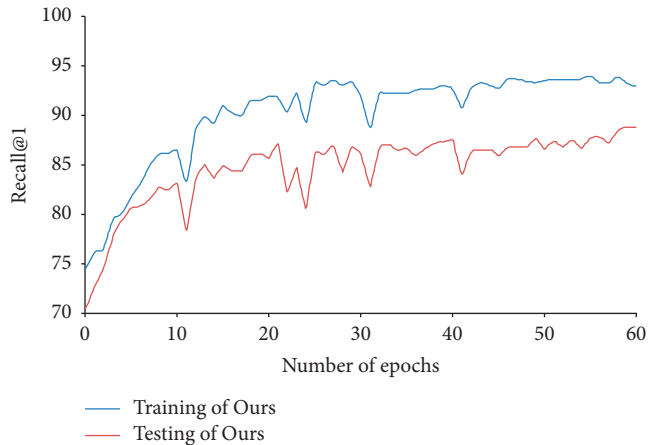


FIGURE 7: Training and testing results (R@1) on Cars-196.

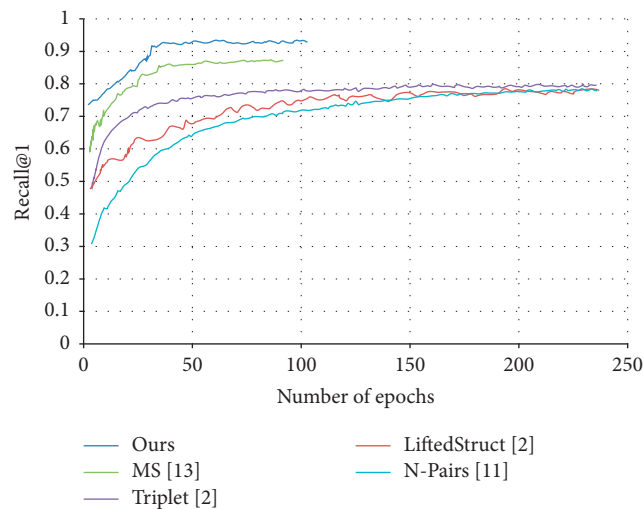


FIGURE 8: Convergence curves of different methods on Cars-196.

Next, the training results and testing results were contrasted on Cars-196. As shown in Figure 7, the training R@1 only had a small gap from the testing R@1. On Cars-196, the R@1 score of the training set was around 93%, only 7% than that on the test set. This clearly shows that our method avoids overfitting.

Figure 8 shows the convergence curves of our method and several state-of-the-art methods on Cars-196. In the first 40 epochs, our method reached the performance of the state of the arts and converged faster than the other methods. However, according to the trend of the curve in Figure 8, that is, number of epochs from 0 to 50, the convergence rate of ours model was not the maximum. However, on the whole, the convergence rate of our method was satisfactory. Except for the MS, the contrastive methods took hundreds of epochs to converge. Thus, the training time of our method was compared with that of the MS. On a single NVIDIA Tesla V100 GPU, the mean running time of our method was 24.36s per epoch on CUB-200-2011 and 40.29s per epoch on Cars-196, while that of the MS was 28.45s and 43.58, respectively.

TABLE 1: Recall@K (%) on CUB-200-2011.

		CUB-200-2011					
R@		1	2	4	8	16	32
Contrastive [2]		26.4	37.7	49.8	62.3	76.4	85.3
Triplet [2]		36.1	48.6	59.3	70.0	80.2	88.4
LiftedStruct [2]		47.2	58.9	70.2	80.2	89.3	93.2
N-Pairs [11]		51.0	63.3	74.3	83.2	—	—
Angular loss [12]		54.7	66.3	76.0	83.9	—	—
HDC [28]		53.6	65.7	77.0	85.6	91.5	95.5
HTL [16]		57.1	68.8	78.7	86.5	92.5	95.5
ABE [9]		60.6	71.5	79.8	87.4	—	—
DREML [10]		63.9	75.0	83.1	89.7	—	—
MS [13]		65.7	77.0	86.3	91.2	95.0	97.3
SoftTriple [29]		65.4	76.4	84.5	90.4	—	—
Ours		69.6	80.3	87.2	93.5	97.6	98.8

Note. DHC and HTL are short for hybrid dilated convolution and hybrid transfer learning, respectively.

TABLE 2: Recall@K (%) on Cars-196.

		Cars-196					
R@		1	2	4	8	16	32
Contrastive [2]		21.7	32.3	46.1	58.9	72.2	83.4
Triplet [2]		39.1	50.4	63.3	74.5	84.1	89.8
LiftedStruct [2]		49.0	60.3	72.1	81.5	89.2	92.8
N-Pairs [11]		71.1	79.7	86.5	91.6	—	—
Angular loss [12]		71.4	81.4	87.5	92.1	—	—
HDC [28]		73.7	83.2	89.5	93.8	96.7	98.4
HTL [16]		81.4	88.0	92.7	95.7	97.4	99.0
ABE [9]		85.2	90.5	94.0	96.1	—	—
DREML [10]		86.0	91.7	95.0	97.2	—	—
MS [13]		84.1	90.4	94.0	96.5	98.0	98.9
SoftTriple [29]		84.5	90.7	94.5	96.9	—	—
Ours		88.8	94.2	96.9	98.4	99.1	99.6

TABLE 3: Recall@K (%) on SOP.

		SOP			
R@		1	10	100	1000
Contrastive [2]		42.0	58.2	73.8	89.1
Triplet [2]		42.1	63.5	82.5	94.8
LiftedStruct [2]		62.1	79.8	91.3	97.4
N-Pairs [11]		67.7	83.8	93.0	97.8
Angular loss [12]		70.9	85.0	93.5	98.0
HDC [28]		69.5	84.4	92.8	97.7
HTL [16]		74.8	88.3	94.8	98.4
ABE [9]		76.3	88.4	94.8	98.2
MS [13]		78.2	90.5	96.0	98.7
SoftTriple [29]		78.3	90.3	95.9	-
Ours		78.7	91.6	96.8	99.3

Finally, the image retrieval efficiency of our method was compared with that of the state-of-the-art methods on CUB-200-2011 and Cars-196, respectively. The comparison results (Tables 1 and 2) show that our method outperformed these methods, including higher-order tuples such as LiftedStruct and N-Pairs, as well as angular loss and ensemble methods such as annotation-based expansion (ABE) and deep randomized ensembles for metric learning (DREML). In particular, on the challenging CUB-200-2011 dataset, our



method led the best-performing state-of-the-art method by a large margin: 4.2% in R@1. On SOP, our method also attained the best performance (Table 3). On all the datasets, our method, with a low feature dimension, performed better than the existing methods, with high feature dimensions.

## 5. Conclusions

This study presents a deep distance metrics ensemble method based on boosting, which generates the final distance metric through iterative training of multiple weak distance metrics. Specifically, the last fully connected layer of the CNN was used to form multiple groups of features. The sample pairs were mapped by the CNN, and the distance between the mapped sample pairs was evaluated by a piecewise linear function. The function was added as a weak learner to the boosting algorithm to generate a strong learner. Then, an alternating optimization method was utilized to optimize the parameters of network and loss function. The effectiveness of our method was demonstrated on three datasets widely used in image retrieval tasks. The future research will further improve our method by cascading more models and combine our method with other loss functions.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported by Scientific Research Project of Xuzhou University of Technology (Grant no. XKY2019107), Science and Technology Project of Construction System, Jiangsu Province, China (Grant no. 2018ZD077), Natural Science Research Project of Jiangsu Province Universities, China (Grant no. 20KJB170023), and Xuzhou Science and Technology Planning Project (Grant no. KC21303).

## References

- [1] B. Liu, R. L. Li, and J. F. Feng, "A brief introduction to deep metric learning," *CAAI Transactions on Intelligent Systems*, vol. 14, no. 6, pp. 1064–1072, 2019.
- [2] H. Oh Song, Y. Xiang, S. Jegelka, and S. Savarese, "Deep metric learning via lifted structured feature embedding," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4004–4012, Las Vegas, Nevada, June 2016.
- [3] Y. Sun, C. Cheng, Y. Zhang et al., "Circle loss: a unified perspective of pair similarity optimization," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 6398–6407, Seattle, WA, USA, June 2020.
- [4] F. Schroff, D. Kalenichenko, and J. Philbin, "Facenet: a unified embedding for face recognition and clustering," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 815–823, Boston, MA, USA, June 2015.
- [5] C. Huang, C. C. Loy, and X. Tang, "Local similarity-aware deep feature embedding," *Advances in Neural Information Processing Systems*, vol. 29, pp. 1262–1270, 2016.
- [6] A. Sanakoyeu, V. Tschernezki, U. Buchler, and B. Ommer, "Divide and conquer the embedding space for metric learning," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 471–480, Long Beach, CA, USA, June 2019.
- [7] H. Xuan, A. Stylianou, and R. Pless, "Improved embeddings with easy positive triplet mining," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 2474–2482, Snowmass Village, CO, USA, March 2020.
- [8] R. Negrel, A. Lechervy, and F. Jurie, "Boosted metric learning for efficient identity-based face retrieval," in *Proceedings of the 26th British Machine Vision Conference*, pp. 139.1–139.12, Swansea, UK, September 2015.
- [9] W. Kim, B. Goyal, K. Chawla, J. Lee, and K. Kwon, "Attention-based ensemble for deep metric learning," in *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 736–751, Munich, Germany, September 2018.
- [10] H. Xuan, R. Souvenir, and R. Pless, "Deep randomized ensembles for metric learning," in *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 723–734, Munich, Germany, September 2018.
- [11] K. Sohn, "Improved deep metric learning with multi-class n-pair loss objective," in *Proceedings of the Advances in Neural Information Processing Systems*, pp. 1857–1865, Barcelona, Spain, December 2016.
- [12] J. Wang, F. Zhou, S. Wen, X. Liu, and Y. Lin, "Deep metric learning with angular loss," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2593–2601, Venice, Italy, October 2017.
- [13] X. Wang, X. Han, W. Huang, D. Dong, and M. R. Scott, "Multi-similarity loss with general pair weighting for deep metric learning," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 5022–5030, Long Beach, CA, USA, June 2019.
- [14] T. T. Do, T. Tran, I. Reid, V. Kumar, T. Hoang, and G. Carneiro, "A theoretically sound upper bound on the triplet loss for improving the efficiency of deep distance metric learning," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 10404–10413, Long Beach, CA, USA, June 2019.
- [15] E. Ustinova and V. Lempitsky, "Learning Deep Embeddings with Histogram Loss," 2016, <https://arxiv.org/abs/1611.00822>.
- [16] W. Ge, W. Huang, D. Dong, and M. R. Scott, "Deep metric learning with hierarchical triplet loss," in *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 269–285, Munich, Germany, September 2018.
- [17] A. Iscen, G. Tolias, Y. Avrithis, and O. Chum, "Mining on manifolds: metric learning without labels," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7642–7651, Salt Lake City, UT, USA, June 2018.
- [18] Y. Zhao, Z. Jin, G. J. Qi, H. Lu, and X. S. Hua, "An adversarial approach to hard triplet generation," in *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 501–517, Munich, Germany, September 2018.
- [19] C. Y. Wu, R. Manmatha, A. J. Smola, and P. Krahenbuhl, "Sampling matters in deep embedding learning," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2840–2848, Venice, Italy, October 2017.

- [20] K. Roth, T. Milbich, and B. Ommer, “Pads: policy-adapted sampling for visual similarity learning,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 6568–6577, Seattle, WA, USA, June 2020.
- [21] J. St. Amand and J. Huan, “Sparse compositional local metric learning,” in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1097–1104, Halifax NS Canada, August 2017.
- [22] J. Bohné, Y. Ying, S. Gentric, and M. Pontil, “Learning local metrics from pairwise similarity data,” *Pattern Recognition*, vol. 75, pp. 315–326, 2018.
- [23] Z. Yan, H. Zhang, R. Piramuthu et al., “Hd-cnn: hierarchical deep convolutional neural networks for large scale visual recognition,” in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2740–2748, Santiago, Chile, December 2015.
- [24] M. Caron, P. Bojanowski, J. Mairal, and A. Joulin, “Unsupervised pretraining of image features on non-curved data,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 2959–2968, Seoul, Korea, November 2019.
- [25] J. H. Friedman, “Greedy function approximation: a gradient boosting machine,” *Annals of Statistics*, pp. 1189–1232, 2001.
- [26] R. E. Schapire and Y. Singer, “Improved boosting algorithms using confidence-rated predictions,” *Machine Learning*, vol. 37, no. 3, pp. 297–336, 1999.
- [27] Z. Zhang, P. Sturgess, S. Sengupta, N. Crook, and P. H. Torr, “Efficient discriminative learning of parametric nearest neighbor classifiers,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2232–2239, Providence, RI, USA, June 2012.
- [28] Y. Yuan, K. Yang, and C. Zhang, “Hard-aware deeply cascaded embedding,” in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 814–823, Venice, Italy, October 2017.
- [29] Q. Qian, L. Shang, B. Sun, J. Hu, H. Li, and R. Jin, “Softtriple loss: deep metric learning without triplet sampling,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 6450–6458, Seoul, South Korea, October 2019.

## Retraction

# Retracted: Local Privacy Protection for Sensitive Areas in Multiface Images

### Computational Intelligence and Neuroscience

Received 3 October 2023; Accepted 3 October 2023; Published 4 October 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] C. Liu, J. Yang, X. Zhang et al., "Local Privacy Protection for Sensitive Areas in Multiface Images," *Computational Intelligence and Neuroscience*, vol. 2022, 15 pages, 2022.

## Research Article

# Local Privacy Protection for Sensitive Areas in Multiface Images

Chao Liu <sup>1,2</sup>, Jing Yang <sup>1</sup>, Xuan Zhang<sup>3</sup>, Yining Zhang <sup>4</sup>, Weinan Zhao <sup>2</sup>,  
Fengjuan Miao<sup>2</sup> and Yukun Shao<sup>2</sup>

<sup>1</sup>College of Computer Science and Technology, Harbin Engineering University, Harbin 150001, China

<sup>2</sup>College of Communication and Electronic Engineering, Qiqihar University, Qiqihar 161000, China

<sup>3</sup>Information Center, First Affiliated Hospital of Qiqihar Medical College, Qiqihar 161000, China

<sup>4</sup>Department of Computer Engineering, Daqing Vocational College, Daqing 163000, China

Correspondence should be addressed to Jing Yang; yangjing@hrbeu.edu.cn

Received 6 January 2022; Revised 26 January 2022; Accepted 28 January 2022; Published 15 March 2022

Academic Editor: Daqing Gong

Copyright © 2022 Chao Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The privacy protection for face images aims to prevent attackers from accurately identifying target persons through face recognition. Inspired by goal-driven reasoning (reverse reasoning), this paper designs a goal-driven algorithm of local privacy protection for sensitive areas in multiface images (face areas) under the interactive framework of face recognition algorithm, regional growth, and differential privacy. The designed algorithm, named privacy protection for sensitive areas (PPSA), is realized in the following manner: Firstly, the multitask cascaded convolutional network (MTCNN) was adopted to recognize the region and landmark of each face. If the landmark overlaps a subgraph divided from the original image, the subgraph will be taken as the seed for regional growth in the face area, following the growth criterion of the fusion similarity measurement mechanism (FSMM). Different from single-face privacy protection, multiface privacy protection needs to deal with an unknown number of faces. Thus, the allocation of the privacy budget  $\epsilon$  directly affects the operation effect of the PPSA algorithm. In our scheme, the total privacy budget  $\epsilon$  is divided into two parts:  $\epsilon_1$  and  $\epsilon_2$ . The former is evenly allocated to each seed, according to the estimated number of faces  $\rho$  contained in the image, while the latter is allocated to the other areas that may consume the privacy budget through dichotomization. Unlike the Laplacian (LAP) algorithm, the noise error of the PPSA algorithm will not change with the image size, for the privacy protection is limited to the face area. The results show that the PPSA algorithm meets the requirements  $\epsilon$ -Differential privacy, and image classification is realized by using different image privacy protection algorithms in different human face databases. The verification results show that the accuracy of the PPSA algorithm is improved by at least 16.1%, the recall rate is improved by at least 2.3%, and  $F1$ -score is improved by at least 15.2%.

## 1. Introduction

In the information age, the importance of data becomes more and more prominent. The value of data is highlighted in various fields in our society. Data-based services are widely applied in different industries. In particular, the application of image data has been developing rapidly. Thanks to the fast development of information technology and multimedia technology, it is now easier to acquire and share digital face images. The users can publish their own photos on social network platforms and other channels. Statistics show that more than 3.2 billion face images are shared by users of major social network platforms around

the world. These digital images usually contain a wealth of personally sensitive information. If the information is collected and analyzed by attackers, unmeasurable losses will occur, in addition to the leak of personal privacy.

The privacy protection of image data often relies on techniques like  $k$ -anonymity, access control, and privacy encryption. Fung et al. [1] and Xiao and Tao [2] proposed the  $k$ -same method based on the anonymization mechanism. The method anonymizes each published gray image, reducing the probability of attackers deriving user identity from the published image to less than  $1/k$ . Li et al. [3] applied access control to restrict user access to social network images. The transfer and visitor volumes of these images are

reduced to protect privacy. This approach, which protects images via access settings, is not a fundamental privacy protection method for images. If an attacker has a certain background knowledge, he/she may bypass the access control and acquire the user images and privacy information. To prevent the communication process from being eavesdropped, Terrovitis et al. [4] implemented homomorphic encryption on gray images. However, every data encryption technique makes some assumptions and designs the corresponding encryption algorithm based on these assumptions. However new it is, the encryption algorithm will soon be breached by attackers. Focusing on anonymized images on Facebook, Sweeney, and Anonymity [5] revealed that attackers can deduce the social security number (SSN) of the person in each anonymized image, based on the unique Friendster feature of Facebook, and identify the sensitive information (e.g., disease and address) from the SSN.

In 2006, Dwork [6] proposed differential probability, which disturbs sensitive data by adding noise to the output. Differential privacy can hide the impact of a single record: whether a record is in the dataset or not, the output probability of the same result will not change significantly. In this way, attackers are hampered from further reasoning. Compared with other privacy protection techniques, the differential probability is superior in that it makes no assumption about the background knowledge of attackers. Dwork further investigated differential privacy in a series of theses [7–11] and proposed its implementation mechanism [12, 13]. McSherry [14] pointed out that the differential privacy algorithm for complex privacy problems needs to satisfy two composition properties: sequence combination and parallel combination. In recent years, differential privacy has been mainly applied to the field of data publication. The differential privacy protection for data publication mainly attempts to ensure the accuracy of published data or query results while satisfying differential privacy. The relevant research focuses on adjusting the publication mechanism and algorithm complexity. The primary research method is a quantitative analysis based on calculation and learning theories.

By the realization environment, the differential privacy protection for data publication can be divided into interactive data publication and noninteractive data publication [12]. On interactive data publication, Roth and Roughgarden [15] presented a median algorithm that can respond to more queries. Hardt and Rothblum [16] developed a pulse wave modulation (PMW) mechanism capable of increasing the number of queries. Gupta et al. [17] proposed a universal iterative dataset generation mechanism. Fan and Xiong [18] designed a novel approach with filtering and adaptive sampling for releasing time series under differential privacy (FAST). Kellaris et al. [19] put forward a flow data publication algorithm with no limit on the number of publications. On noninteractive data publication, Xiao et al. [20] came up with the Privelet algorithm. Xu et al. [21] created the noise first and structure first algorithms. Li et al. [22] proposed the matrix mechanism. Li et al. [23] put forward a data- and workload-aware algorithm (DAWA) algorithm.

Owing to the complexity of image data, the current research is still in the exploratory stage concerning the differential privacy protection for sensitive information in images.

Images are often represented as matrices in the real number field. Any pixel in an image can be mapped to a value at the corresponding location in a two-dimensional (2D) matrix. To satisfy  $\epsilon$ -differential privacy, the most direct approach is to add a Laplace noise to all the values in the matrix. However, the disturbed image may be over distorted and become useless. Fourier transform and wavelet transform are common ways to compress images. Zhang et al. [24] proposed an image compression method based on discrete Fourier transform (DFT), which adds a unique Laplace noise to each target image. Despite suppressing the noise error, their approach introduces the reconstruction error to image compression. Considering the uncorrelation between the values in the image matrix, Liu et al. [25] converted the image gray matrix into a one-dimensional (1D) ordered data flow and modeled the data flow with the sliding window model. To protect the privacy of images, the privacy budget was allocated dynamically based on the similarity between the data of adjacent sliding windows. This highly available strategy is restricted to the 1D space. Liu et al. [26] utilized regional growth to expand the comparison between adjacent subgraphs to the 2D space and further optimized the privacy protection for face images.

The noises of the above methods cover the entire image. For face images, the sensitive information of the face only concentrates in the face area and even in some specific areas. The exposure of other nonsensitive areas will not lead to the leak of privacy. Therefore, this paper proposes a local privacy protection method for the sensitive areas in multiface images.

## 2. Background

### 2.1. Differential Privacy

*Definition 1.* Adjacent datasets of the face image

For a given image  $X$ , the gray matrix  $X_{mn}$  can be obtained by normalizing the image. Then, there exists

$$X|X_{mn} = \begin{bmatrix} x_{11}, x_{12}, & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1}, x_{m1}, & \cdots & x_{mn} \end{bmatrix}, \text{ where } x_{ij} \text{ in matrix } X_{mn}$$

represents the gray of the corresponding element. If there exists an  $X'$  with only one element difference from  $X$ ,  $|X - X'| = x_{ij}$  ( $1 \leq i \leq m, 1 \leq j \leq n$ ), then  $X$  and  $X'$  are adjacent datasets.

*Definition 2.* Differential privacy

For a given random algorithm  $M$  of image data publication, with the output range of  $\text{Range}(M)$ , the algorithm can provide  $\epsilon$ -differential privacy, if its arbitrary outputs on two adjacent gray images  $X$  and  $X'$  satisfy the following:

$$\Pr[M(X) \in S] \leq \exp(\epsilon) \times \Pr[M(X') \in S], \quad (1)$$

where  $\epsilon$  is typically a small positive number that balances privacy with accuracy. If  $\epsilon$  is small, the privacy is high and

accuracy is low. The inverse is also true. Normally,  $\epsilon$  is selected by the user by executing a privacy policy. When the adjacent datasets vary by only one record, the algorithm satisfies  $\epsilon$ -differential privacy. When the adjacent datasets vary by  $k$  records, the algorithm satisfies  $k\epsilon$ -differential privacy.

*Definition 3.* Global sensitivity

Let  $Q$  be a random query function meeting  $Q: D \rightarrow R^n$ . Then, the global sensitivity of  $Q$  can be expressed as follows:

$$\Delta Q_{GS} = \max_{X, X'} \|Q(D) - Q(D')\|_p. \quad (2)$$

**Theorem 1.** Laplace mechanism

Let  $Q$  be a query series of the length  $d$ . The random algorithm  $M$  receives database  $D$  and outputs the following vector that satisfies  $\epsilon$ -differential privacy:

$$M(D) = Q(D) + \langle \text{Lap}_1\left(\frac{\Delta Q_{LS}}{\epsilon}, \dots, \text{Lap}_d\left(\frac{\Delta Q_{LS}}{\epsilon}\right) \right). \quad (3)$$

As the most common noise addition mechanism, the Laplace mechanism disturbs the real output by adding the noise generated by Laplace distribution, thereby achieving differential privacy. The probability density function (PDF) of its noise distribution satisfies  $f(x|\mu, b) = 1/2be^{|x-\mu|/b}$ .

*Property 1.* Differential privacy-serial combination property

For a given dataset  $X$  and a set of differential privacy algorithms  $M_1(X), M_2(X), \dots, M_m(X)$  related to  $X$ , if the algorithm  $M_i(D)$  satisfies  $\epsilon_i$ -differential privacy, and the random processes of any two algorithms are independent of each other, then the algorithm combined from these algorithms satisfies  $\sum_{i=1}^m \epsilon_i$ -differential privacy.

## 2.2. Multitask Cascaded Convolutional Network (MTCNN).

With the development of CNN [27], Sun et al. [28] suggested that the CNN can be applied to localization face landmarks by virtue of its strong feature extraction ability. A three-layer deep CNN (DCNN) was designed to solve a thorny problem in landmark extraction: the inability to obtain the global optimal solution, due to the improper setting of initial values. Zhang et al. introduced multitask learning to face landmark localization and proposed a task-driven DCNN for face landmark localization (TCDCN). The TCDCN, a multitask learning model of four subtasks, has a smaller time complexity than traditional CNN [29]. Two years later, Zhang et al. developed the MTCNN, which effectively integrates face area detection with face landmark localization. The three cascading neural networks in MTCNN are responsible for face classification, bounding box regression, and key point localization, respectively [30]. Wu et al. [31] presented a tweaked CNN (TCNN), which relies on a mixed Gaussian model to cluster the features on different layers, and concluded that a deeper network layer can more accurately mirror face landmarks. In addition, many other methods have been adopted to locate face landmarks,

namely, principal component analysis (PCA) [32], support vector machine [33], Bayesian probabilistic network (BPN) [34], dynamic link architecture (DLA) [35], and Gabor wavelet network (GWN) [36].

The MTCNN is a multitask parallel face recognition method based on deep learning. This method has been widely recognized in the industry because it operates rapidly, and its accuracy meets the general requirements of face detection. The multitask parallel capacity of the MTCNN algorithm manifests as the simultaneous detection of multiple faces (detecting whether an image contains faces and finding the face locations in the image, as shown in Figure 1) and the localization of face landmarks (locating the five landmarks, namely, two eyes, nose tip, and mouth corners, as shown in Figure 2).

The MTCNN algorithm can be realized as a three-stage cascaded CNN. The first stage is called a proposal network (P-Net), which mainly obtains the candidate windows for the face area and the regression vector of the bounding box, carries out regression with the bounding box, calibrates the candidate windows, and merges highly overlapped candidate boxes through nonmaximum suppression (NMS). The structure of the P-Net is shown in Figure 3.

The second stage is called a refine network (R-Net), which eliminates the false positive areas through bounding box regression and NMS. That is, lots of nonface windows are denied by a more complex CNN in order to refine the face window. The structure of the R-Net is shown in Figure 4.

The third stage is called an output network (O-Net). With one more convolutional layer than the R-Net, the O-Net outputs more refined results. The functions of the O-Net are the same as those of the R-Net. But the O-Net carries out more supervision of the face area and output five landmarks. The structure of the O-Net is shown in Figure 5.

Each stage of the MTCNN is a multitask network, responsible for handling tasks like face/nonface judgement, face box regression, and landmark localization. The face/nonface judgement adopts the cross-entropy loss function:

$$L_i^{\text{det}} = -\left(y_i^{\text{det}} \log(p_i) + (1 - y_i^{\text{det}})(1 - \log(p_i))\right), \quad (4)$$

where  $p_i$  is the probability for the image to contain faces;  $y_i^{\text{det}} \in \{0, 1\}$  is the true background label.

The regression loss of face box can be calculated by Euclidean distance:

$$L_i^{\text{box}} = \|\hat{y}_i^{\text{box}} - y_i^{\text{box}}\|_2^2, \quad (5)$$

where  $\hat{y}$  is the background coordinates predicted by the network;  $y$  is the true background coordinates;  $y_i^{\text{box}} \in \mathbb{R}^4$  ( $\mathbb{R}^4$  is a 4-tuple composed of  $x$  of the upper left corner,  $y$  of the upper left corner, length, and width).

Landmark localization is calculated similarly as face box regression. The Euclidean distance between the predicted landmark location and the true landmark location is calculated and minimized:

$$L_i^{\text{landmark}} = \|\hat{y}_i^{\text{landmark}} - y_i^{\text{landmark}}\|_2^2, \quad (6)$$

where  $y_i^{\text{landmark}} \in \mathbb{R}^{10}$  is a 10-tuple, for five landmarks are being localized.



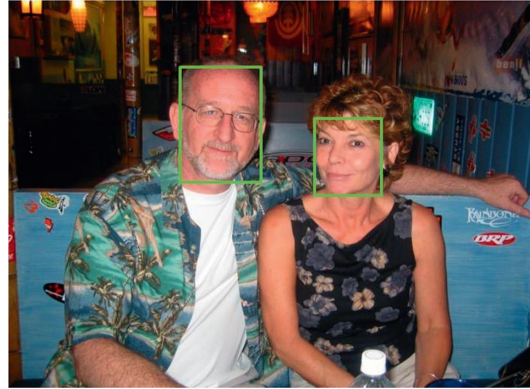


FIGURE 1: Face detection.



FIGURE 2: Landmark localization.

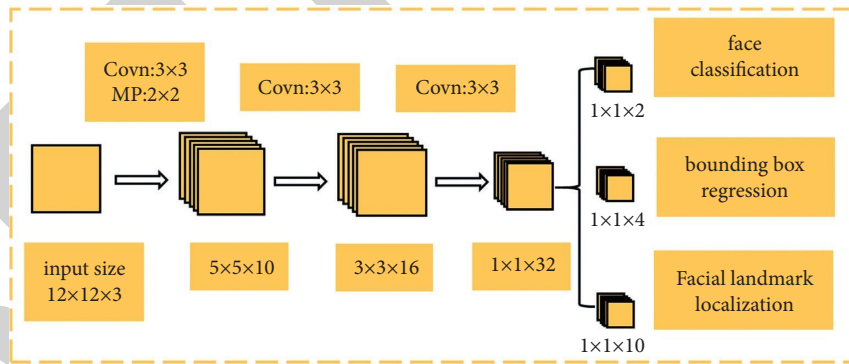


FIGURE 3: Structure of P-net.

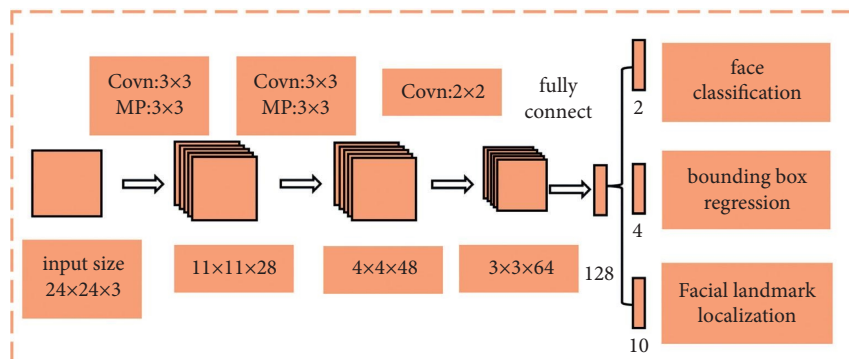


FIGURE 4: Structure of R-net.

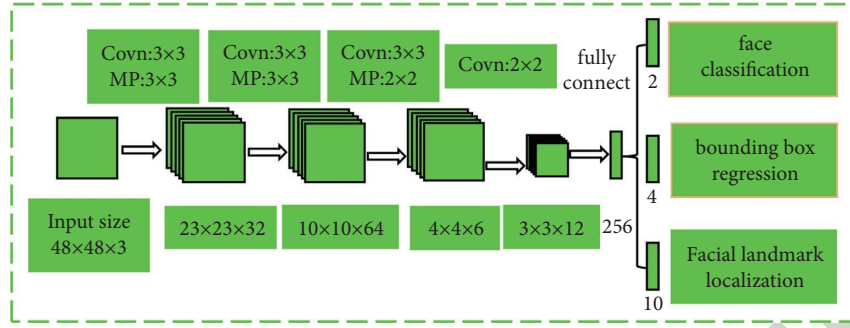


FIGURE 5: Structure of O-net.

The simplest function of the training process can be expressed as follows:

$$\min \sum_{i=1}^N \sum_{j \in \{\text{det}, \text{box}, \text{landmark}\}} \alpha_j \beta_i^j L_i^j, \quad (7)$$

where  $N$  is the number of training samples;  $\alpha_j$  is the importance of a task;  $\beta_i^j \in \{0, 1\}$  is a sample label. In the P-Net and R-Net,  $\alpha_{\text{det}} = 1$ ,  $\alpha_{\text{box}} = 0.5$ , and  $\alpha_{\text{landmark}} = 0.5$ . In the O-Net,  $\alpha_{\text{det}} = 1$ ,  $\alpha_{\text{box}} = 0.5$ , and  $\alpha_{\text{landmark}} = 1$ .

### 3. Methodology

**3.1. Laplacian (LAP) Algorithm.** This paper proposes the LAP algorithm based on the Laplace mechanism. In the LAP algorithm, every element  $x_{ij}$  in the gray matrix  $X_{m \times n}$  is regarded as an independent individual. This division method is the basis for applying the interactive mechanism to privacy protection. Each  $x_{ij}$  ( $1 \leq i \leq m$ ,  $1 \leq j \leq n$ ) consumes a privacy budget of  $\epsilon / (m \times n)$ . The LAP algorithm can be realized in the steps in Algorithm 1.

Compared with the original image  $X$ , the privacy protected image  $X'$  contains an additive noise of  $2m \times n \times (\Delta Q \times m \times n / \epsilon)^2$ . Admittedly, the LAP algorithm satisfies  $\epsilon$ -differential privacy. However, the noise results show that the LAP algorithm will have a huge error if it is applied to protect the privacy of an excessively large image. If so, the noisy image will be of low availability.

**3.2. Fusion Similarity Measurement Mechanism (FSMM).** This paper divides a face image  $X_{m \times n}$  into multiple non-intersecting subgraphs  $T_{ij}$  ( $1 \leq i \leq I$ ,  $1 \leq j \leq J$ ), each of which contains multiple pixels:

$$X_{m \times n} = \begin{bmatrix} x_{11}, x_{12}, & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1}, x_{m1}, & \cdots & x_{mn} \end{bmatrix} = \begin{bmatrix} T_{11}, T_{12}, & \cdots & T_{1J} \\ \vdots & \ddots & \vdots \\ T_{I1}, T_{I1}, & \cdots & T_{IJ} \end{bmatrix}. \quad (8)$$

Through the division, the nonintersecting subgraphs  $T_{ij}$  carry as much information of the original image as possible.

The key to regional growth is to determine the criterion. Due to the natural complexity of face images, the criterion should be designed by judging the various features of the original image. The traditional criterion for regional growth only focuses on the difference between gray values. The gray value of a single pixel cannot provide rich information about the image. By contrast, subgraphs as the basic units of regional growth can retain the luminance, contrast, structure, color, texture, and spatial distribution features of the image. This paper presents a new regional growth criterion called FSMM, which further improves the accuracy of area merging with subgraphs as the basic units. The FSMM can be realized in the steps in Algorithm 2.

As its name suggests, the new regional growth criterion relies on the FSMM, a brand-new similarity measure between images. In rows 3-4 of the algorithm, parameter  $\sigma$ , a very small positive number, appears to ensure the denominator is nonzero. In rows 7-9,  $u_X$  and  $u_Y$  are, respectively, the mean values of images  $X$  and  $Y$ , reflecting the luminance features of the images;  $\sigma_X$  and  $\sigma_Y$  are, respectively, the variances of images  $X$  and  $Y$ , reflecting the contrast features;  $C_1, C_2$  and  $C_3$  are very small positive numbers that ensure the denominator is nonzero. In Column 10,  $\alpha, \beta$ , and  $\gamma$  are parameters adjusting the proportion of different eigenvalues: if

$\alpha = \beta = \gamma = 1$ , then  $C_1 = (K_1 L)^2$ ,  $C_2 = (K_2 L)^2$ , and  $C_3 = C_2/2$ , with  $K_1 \ll 1$ ,  $K_2 \ll 1$ , and  $L$  being the dynamic range of the image. The final calculation method for FSMM( $X, Y$ ) can be expressed as follows:

$$\text{FSMM}(X, Y) = \frac{\phi \times (d(X, Y) / \rho(X, Y) + \sigma) \times (2u_X u_Y + C_1) \times (2\sigma_X \sigma_Y + C_2)}{(u_X^2 + u_Y^2 + C_1) \times (\sigma_X^2 + \sigma_Y^2 + C_2)}. \quad (9)$$

Input: original image  $X$ , privacy budget  $\varepsilon$ , parameters  $m$  and  $n$ , subgraph similarity expectation  $Th$   
 Output: Image  $X'$  satisfying differential privacy

- (1) Read the original image  $X$ , convert the image into gray matrix and store it in matrix  $X_{m \times n}$
- (2) for  $i = 1$  to  $m$
- (3)   for  $i = 1$  to  $n$
- (4)     $X'_{(i,j)} = X_{(i,j)} + \text{lap}(\Delta Q \times m \times n/\varepsilon)$
- (5)   end for
- (6)   end for
- (7) Output privacy protected picture  $X'$

ALGORITHM 1: LAP.

Input: image  $X$ , image  $Y$   
 Output: graph similarity  $\theta$

- (1) Convert the image to a grayscale image,  $X' = \begin{bmatrix} X_1 & \dots & \dots \\ \dots & \ddots & \dots \\ \dots & \dots & X_c \end{bmatrix}, Y' = \begin{bmatrix} Y_1 & \dots & \dots \\ \dots & \ddots & \dots \\ \dots & \dots & Y_c \end{bmatrix}$
- (2) Calculate the Euclidean distance between  $X'$  and  $Y'$ ,  $\theta = \sqrt{\sum_{i=1}^n (X_i - Y_i)^2}$
- (3) Calculate the Jaccard similarity index between  $X'$  and  $Y'$ ,  $\rho(X', Y') = |X' \cap Y'| / |X' \cup Y'|$
- (4)  $\theta = \theta / \rho(X', Y') + \sigma$
- (5) Combined with Hamming distance and perceptual hash algorithm  $\phi$ ,  $\phi = \begin{cases} \theta \times (\rho(X', Y') + \sigma) / d(X', Y') \mu(X', Y') < 0.078 \times c \\ 1, 0.078 \times c \leq \mu(X', Y') \leq 0.156 \times c \\ (1/\lg) \times (m \times n - \mu(X', Y')), \mu(X', Y') > 0.156 \times c \end{cases}$
- (6)  $\theta = \phi \times \theta$
- (7) Calculate the luminance relationship between  $X'$  and  $Y'$ ,  $l(X', Y') = (2u_{X'}u_{Y'} + C_1) / (u_{X'}^2 + u_{Y'}^2 + C_1)$
- (8) Calculate the contrast ratio relationship between  $X'$  and  $Y'$ ,  $c(X', Y') = (2\sigma_{X'}\sigma_{Y'} + C_2) / (\sigma_{X'}^2 + \sigma_{Y'}^2 + C_2)$
- (9) Calculate the structural relationship between  $X'$  and  $Y'$ ,  $s(X', Y') = (\sigma_{X'Y'} + C_3) / (\sigma_{X'}\sigma_{Y'} + C_3)$
- (10)  $\text{SSIM}(X', Y') = [l(X', Y')]^a \cdot [c(X', Y')]^b \cdot [s(X', Y')]^y$
- (11)  $\theta = \text{SSIM}(X', Y') \times \theta$
- (12) Output  $\theta$

ALGORITHM 2: FSMM.

### 3.3. Privacy Protection for Sensitive Areas (PPSA) Algorithm.

To achieve privacy protection of sensitive information (face area) in multiface images and reduce noise impact on the privacy protected images, this paper proposes a local privacy protection method for face images, which combines face recognition, regional growth, and differential privacy. The proposed algorithm is called the PPSA.

Inspired by goal-driven reasoning (reverse reasoning), the PPSA was designed with goal as the starting point. The privacy protection for face images aims to prevent attackers from recognizing the target persons through face detection. Therefore, the first step of PPSA design is to understand how attackers identify faces. Next, a privacy protection method should be provided for face images, targeting the face identification technique of attackers. In this way, the designed privacy protection algorithm can effectively curb privacy leak and mitigate the effects of noise on the availability of the privacy protected image.

The design of the PPSA needs to solve three key problems: (1) locate the sensitive area in the multiface image and recognize the landmarks in the area; (2) formulate the criterion for regional growth; (3) allocate the privacy budget reasonably to areas requiring different levels of protection.

The PPSA algorithm is designed to protect the sensitive information (face information) in multiface images in batches. Different from single-face privacy protection,

multiface privacy protection needs to deal with an unknown number of faces. Thus, the allocation of the privacy budget is crucial to PPSA design. The privacy protection should be tailored to areas with different protection requirements. In our scheme, the total privacy budget  $\varepsilon$  is divided into two parts:  $\varepsilon_1$  and  $\varepsilon_2$ . The  $\varepsilon_1$  is evenly allocated to each seed  $\text{Seed}_n$  (any subgraph containing a landmark is a seed), according to the estimated number of faces  $\rho$  contained in the image. During regional growth,  $\text{Seed}'_n$  (noisy seed) replaces the  $T_{(i,j)}$  (adjacent subgraph) meeting the growth criterion, thereby protecting the privacy of  $T_{(i,j)}$ . Note that the privacy protection of  $T_{(i,j)}$  does not consume any privacy budget. Meanwhile, the  $\varepsilon_2$  is allocated through dichotomization. This part of the privacy budget is consumed by the  $\text{Seed}_n$  outside the estimated number of faces  $\rho$ , the  $T_{(i,j)}$  failing to meet the growth criterion, and the  $\tilde{T}_{(i,j)}$  belonging to the sensitive area but not involved in regional growth. The PPSA algorithm can be realized in the steps in Algorithm 3.

The image preprocessing, privacy budget allocation, and regional growth are described in rows 1–6, rows 7–9, and rows 10–29, respectively. In row 30,  $\tilde{T}_{(i,j)}$  is the addition of Laplace noise to an unprotected subgraph in  $S'$ . In row 18, the PPSA limits the range of regional growth: if a  $T_{(i,j)}$  meeting the growth criterion but lying beyond the coverage of the current  $S'_i$  or the growth range of the current  $\text{Seed}_n$ , it will not be able to complete this area merging; the growth

Input: Original image  $X$ , privacy budget  $\varepsilon$ , estimated number of faces  $\rho$ , preset parameters  $\beta$ , estimate the number of faces in the image  $\omega$ , subgraph similarity expectation  $Th$

Output: Image  $X'$  satisfying differential privacy.

- (1) Read the original image  $X$ , convert the image into gray matrix and store it in matrix  $X_{m \times n}$
- (2) MTCNN is used to extract the face region  $S = (S_1, S_2, \dots, S_\omega)$  and face feature point  $K_S = (K_1, K_2, \dots, K_{5 \times \omega})$  in face image  $X$
- (3) Set the gray matrix  $X$  according to the preset parameters  $\beta$  It is divided into subgraph sets  $T_{(I,J)}$  with the same structure
- (4) If  $K_S$  belongs to a subgraph  $T_{(i,j)}$ , set  $T_{(i,j)}$  as  $\text{Seed}_n$
- (5) Find all backup  $\text{Seed}_N = (\text{Seed}_1, \text{Seed}_2, \dots, \text{Seed}_{5 \times \omega})$
- (6) Adjust the face image area  $S$  as  $S'$  according to the sub image size
- (7)  $\varepsilon = \varepsilon_1 + \varepsilon_2$
- (8)  $\varepsilon_1^{\text{left}} = \varepsilon_1$
- (9)  $\varepsilon_2^{\text{left}} = \varepsilon_2/2$
- (10) Set  $\text{Seed} = 0$ ,  $\widetilde{\text{Seed}}$  used to record whether this seed has been used during regional growth
- (11) Create a linked **list** to record the status of the current seed region merging process
- (12) while ( $\text{Seed} \leq 5 \times \omega$ )
- (13) if ( $\varepsilon_1^{\text{left}} > 0$ )
- (14) if ( $\text{list\_head} == \text{NULL}$ )
- (15) Pick unmarked  $\text{Seed}_n$  adds noise to it as  $\text{Seed}'_n$ ,  $\text{Seed}'_n = \text{Seed}_n + \text{lap}(\Delta Q \times 5 \times \rho / \varepsilon_1)$
- (16)  $\varepsilon_1^{\text{left}} = \varepsilon_1^{\text{left}} - \varepsilon_1 / 5\rho$
- (17) Attempt to merge adjacent  $T_{(i,j)}$  start regional growth
- (18) if ( $T_{(i,j)} \notin S' \parallel \text{Euclidean\_Distance}(T_{(i,j)}, \text{Seed}_n) \geq \text{Min}(\text{Euclidean\_Distance}(\text{Seed}_n, \text{Seed}_m) / 2)$ )
- (19) brake
- (20) else
- (21) if ( $\text{FSMM}(\text{Seed}'_n, T_{(i,j)}) \leq Th$ )
- (22)  $T_{(i,j)} = \text{Seed}'_n$
- (23) else
- (24)  $T'_{(i,j)} = T_{(i,j)} + \text{lap}(\Delta Q / \varepsilon_2^{\text{left}})$
- (25)  $\varepsilon_2^{\text{left}} = \varepsilon_2^{\text{left}} / 2$
- (26)  $\text{Seed}_n$  is set to marked,  $\widetilde{\text{Seed}} = \widetilde{\text{Seed}} + 1$
- (27) Release all data in the list, and the consolidation of this subregion is completed
- (28) else if ( $\varepsilon_1^{\text{left}} = 0$ )
- (29) Execute line 15–line 28, line 16  $\Delta Q \times 5 \times \rho / \varepsilon_1$  adjusted to  $\Delta Q / \varepsilon_2^{\text{left}}$ , line 17 is adjusted to  $\varepsilon_2^{\text{left}} = \varepsilon_2^{\text{left}} / 2$
- (30)  $\widehat{T}_{(i,j)} = \widehat{T}_{(i,j)} + \text{lap}(\Delta Q / \varepsilon_2^{\text{left}})$

ALGORITHM 3: PPSA.

range of  $\text{Seed}_n$  will be limited to half of the Euclidean distance to the nearest adjacent  $\text{Seed}_m$ . If  $\omega \leq \rho$ , only rows 10–27 will be executed for regional growth. If  $\omega > \rho$ , all the rows from 10–29 will be executed for regional growth.

**Theorem 2.** *The PPSA consumes a privacy budget smaller than  $\varepsilon$  and satisfies  $\varepsilon$ -differential privacy.*

*Proof.* In the PPSA, the total privacy budget  $\varepsilon$  is divided into two parts. Among them,  $\varepsilon_1$  is used to add noise to seeds. If  $\omega < \rho$ , there exists  $\varepsilon_1^{\text{left}} > 0$  after the end of the algorithm; if  $\omega \geq \rho$ , there exists  $\varepsilon_1^{\text{left}} = 0$  after the end of the algorithm. Meanwhile,  $\varepsilon_2$  is allocated by dichotomy:

$$\varepsilon_2 \times \left( \frac{1}{2^1} + \frac{1}{2^2} + \frac{1}{2^3} + \dots + \frac{1}{2^N} \right) < \varepsilon_2 \quad (N \rightarrow +\infty). \quad (10)$$

Thus, there exists the following:

$$\varepsilon - (\varepsilon_1^{\text{left}} + \varepsilon_2^{\text{left}}) > 0. \quad (11)$$

Therefore, the PPSA consumes a privacy budget smaller than  $\varepsilon$ . According to Property 1, the PPSA satisfies  $\varepsilon$ -differential privacy.

Q.E.D.

**Theorem 3.** *The error of the PPSA is no greater than that of the LAP, i.e.,*

$$\text{Error (DPLP)} < \text{Error (LAP)}. \quad (12)$$

*Proof.* In the PPSA, the privacy protected image contains four kinds of subgraphs, namely, noise-free subgraph  $P_{(i,j)}$ , seed  $\text{Seed}_n$ , subgraph merged in regional growth  $T_{(i,j)}$ , and subgraph belonging to a sensitive area but not involved in regional growth  $\widehat{T}_{(i,j)}$ .

(1) For  $P_{(i,j)}$ :

$$P'_{(i,j)\text{-PPSA}} = 0$$

$$P'_{(i,j)\text{-LAP}} = P_{(i,j)} + \text{lap}\left(\frac{\Delta Q \times I \times J}{\varepsilon}\right). \quad (13)$$

Subtracting the two formulas above, we have the following:  $\text{Error (PPSA}(P_{(i,j)})) < \text{Error (LAP}(P_{(i,j)}))$ .

(2) For  $\widehat{T}_{(i,j)}$ :

$$\begin{aligned}\hat{T}'_{(i,j)\text{-LAP}} &= \hat{T}_{(i,j)} + \text{lap}\left(\frac{\Delta Q \times I \times J}{\varepsilon}\right) \\ \hat{T}'_{(i,j)\text{-PPSA}} &= \hat{T}_{(i,j)} + \text{lap}\left(\frac{\Delta Q \times 2^M}{\varepsilon_2}\right).\end{aligned}\quad (14)$$

If  $2^M \leq I \times J$ , there exists  $\text{Error}(\text{PPSA}(P_{(i,j)})) < \text{Error}(\text{LAP}(P_{(i,j)}))$ .

If  $2^M > I \times J$ , there exists  $\text{Error}(\text{PPSA}(P_{(i,j)})) \approx \text{Error}(\text{LAP}(P_{(i,j)}))$ .

The validity of  $\text{Error}(\text{PPSA}(P_{(i,j)})) \approx \text{Error}(\text{LAP}(P_{(i,j)}))$  depends on the uniqueness of the gray image. In the gray image matrix, the gray value falls between 0 and 255. Under the Laplace mechanism, the smaller the  $\varepsilon$ , the greater the noise. For any other type of data, the noise effect will be huge, if  $\varepsilon$  is sufficiently small. For gray value, however, the fluctuation induced by noise is limited. Take  $\hat{T}'_{(i,j)\text{-PPSA}} = \hat{T}_{(i,j)} + \text{lap}(\Delta Q \times 2^M / \varepsilon_2)$  for example. When  $\text{lap}(\Delta Q \times 2^M / \varepsilon_2) \rightarrow +\infty$  时,  $\hat{T}'_{(i,j)\text{-PPSA}}$  value in the gray matrix will become 255. When  $\text{lap}(\Delta Q \times 2^M / \varepsilon_2) \rightarrow -\infty$ , that value will change to 0. Thus,  $\text{Error}(\text{PPSA}(P_{(i,j)})) \approx \text{Error}(\text{LAP}(P_{(i,j)}))$  holds, as long as  $I \times J$  is sufficiently great.

(3) For  $\text{Seed}_n$ :

There are two possible scenarios:  $\omega \leq \rho$  and  $\omega > \rho$ .

When  $\omega \leq \rho$ ,  $\text{Seed}'_{n\text{-DPLP}}$  only consumes  $\varepsilon_1$ . Then, there exists the following:

$$\begin{aligned}\text{Seed}'_{n\text{-PPSA}} &= \text{Seed}_n + \text{lap}\left(\frac{\Delta Q \times 5 \times \rho}{\varepsilon_1}\right) \\ \text{Seed}'_{n\text{-LAP}} &= \text{Seed}_n + \text{lap}\left(\frac{\Delta Q \times I \times J}{\varepsilon}\right).\end{aligned}\quad (15)$$

If  $5 \times \rho \ll I \times J$ , the numerical gap between  $\varepsilon_1$  and  $\varepsilon$  is so small as to be negligible. Subtracting the above two formulas, we have the following:  $\text{Error}(\text{PPSA}(\text{Seed}_n)) < \text{Error}(\text{LAP}(\text{Seed}_n))$ .

When  $\omega > \rho$ , the first  $\rho$  face areas consume  $\varepsilon_1$  until no budget in this part is left. In this case, there exists  $\text{Error}(\text{PPSA}(\text{Seed}_n)) < \text{Error}(\text{LAP}(\text{Seed}_n))$ . For the lack of space, the proof is omitted.

In the following  $\omega - \rho$  face images,  $\text{Seed}'_{n\text{-DPLP}}$  needs to consume  $\varepsilon_2$ . According to the proof of  $\hat{T}'_{(i,j)}$ , there exists  $\text{Error}(\text{PPSA}(\text{Seed}_n)) \approx \text{Error}(\text{LAP}(\text{Seed}_n))$  during the consumption of  $\varepsilon_2$ . Thus, it can be deduced that, in any scenario, there exists the following:

$$\begin{aligned}\text{Error}(\text{PPSA}(\text{Seed}_n))_{\varepsilon_1} + \text{Error}(\text{PPSA}(\text{Seed}_n))_{\varepsilon_2} \\ < \text{Error}(\text{LAP}(\text{Seed}_n))_{\varepsilon}.\end{aligned}\quad (16)$$

(4) For  $T_{(i,j)}$ :

There are also two possible scenarios:  $\omega \leq \rho$  and  $\omega > \rho$ . When  $\omega \leq \rho$ , the proof is as follows:

$$\begin{aligned}\text{Error}(\text{PPSA}(T_{(i,j)})) &= T'_{(i,j)\text{-PPSA}} = \text{Seed}_n + \text{lap}\left(\frac{\Delta Q \times 5 \times \rho}{\varepsilon_1}\right) - T_{(i,j)} \\ \text{Error}(\text{LAP}(T_{(i,j)})) &= T'_{(i,j)\text{-LAP}} = \text{lap}\left(\frac{\Delta Q \times I \times J}{\varepsilon}\right).\end{aligned}\quad (17)$$

Subtracting the above two formulas, we have the following:

$$\begin{aligned}\text{Error}(\text{PPSA}(T_{(i,j)})) - \text{Error}(\text{LAP}(T_{(i,j)})) \\ &= \text{Seed}_n + \text{lap}\left(\frac{\Delta Q \times 5 \times \rho}{\varepsilon_1}\right) - T_{(i,j)} - \text{lap}\left(\frac{\Delta Q \times I \times J}{\varepsilon}\right) \\ &= \left(\text{Seed}_n + \text{lap}\left(\frac{\Delta Q \times 5 \times \rho}{\varepsilon_1}\right)\right) - \left(T_{(i,j)} + \text{lap}\left(\frac{\Delta Q \times I \times J}{\varepsilon}\right)\right) \\ &= \text{Seed}'_{n\text{-PPSA}} - T'_{(i,j)\text{-LAP}}.\end{aligned}\quad (18)$$

For a gray image, when the privacy budget is sufficiently small under the Laplace mechanism,  $\text{Error}(T'_{(i,j)\text{-LAP}})$  is equivalent to  $\text{Error}(\text{Seed}'_{n\text{-LAP}})$ . Thus, we have the following:

$$\begin{aligned}\text{Seed}'_{n\text{-PPSA}} - T'_{(i,j)\text{-LAP}} \\ &= \text{Seed}'_{n\text{-PPSA}} - \text{Seed}'_{n\text{-LAP}} \\ &= \text{Seed}_n + \text{lap}\left(\frac{\Delta Q \times 5 \times \rho}{\varepsilon_1}\right) - \text{Seed}_n - \text{lap}\left(\frac{\Delta Q \times I \times J}{\varepsilon}\right) \\ &= \text{lap}\left(\frac{\Delta Q \times 5 \times \rho}{\varepsilon_1}\right) - \text{lap}\left(\frac{\Delta Q \times I \times J}{\varepsilon}\right) < 0.\end{aligned}\quad (19)$$

Thus,  $\text{Error}(\text{PPSA}(T_{(i,j)})) < \text{Error}(\text{LAP}(T_{(i,j)}))$ .

Similarly, when  $\omega > \rho$ , there exists  $\text{Error}(\text{PPSA}(T_{(i,j)})) \approx \text{Error}(\text{LAP}(T_{(i,j)}))$ . For the lack of space, the proof is omitted.

Q.E.D.





FIGURE 6: Original image (RGB).



FIGURE 7: Gray image.

FIGURE 8:  $S$  and  $K_S$ .

## 4. Experiments and Results Analysis

**4.1. Experiments.** To demonstrate its feasibility, the PPSA was tested on a  $750 * 1020$  image containing 8 faces from the WIDER FACE dataset. The original image was divided into 7,650 subgraphs of the size  $10 * 10$ . During the implementation of the PPSA, the original image (Figure 6) was first transformed into a gray image (Figure 7). Then, the MTCNN algorithm was called to recognize the face area  $S$  and landmark  $K_S$  (Figure 8). To meet the requirements of the

PPSA on regional growth, the face image was divided into multiple equal size  $T_{(I,J)}$  (Figure 9), using the preset parameters. The subgraph area (Figure 10) overlapping the landmark location was taken as the  $Seed_N$  (Figure 11) of regional growth. However, the  $S$  detected by the MTCNN is not compatible with the size  $fT_{(I,J)}$  (Figure 12). To solve the problem,  $S$  was resized to  $S'$  (Figure 13), according to the size of  $T_{(I,J)}$ .

Figure 14 provides the results of the LPA. Figure 15 presents the PPSA results under the same privacy budget.



FIGURE 9: Rendering of  $T(i, j)$ .FIGURE 10: Rendering of  $K_S$  and  $T(i, j)$ .FIGURE 11: Rendering of  $Seed_N$ .

Observations show that the PPSA operates within the face area while the LAP acts across the image. Besides, within the scope of the regional growth algorithm, the PPSA retained the recognizability of landmarks. Although not excluding the possibility of privacy leaks, this strikes a balance between the privacy and availability of the privacy protected multiface image. In addition, the relationship between the number of faces  $\rho$  estimated by the PPSA and the true number of faces  $\omega$  determines how much privacy budget can be obtained by a seed. The allocation of the privacy budget

will directly affect the error of the protected image. Figures 16 and 17 present the local results at different relationships between  $\rho$  and  $\omega$ .

**4.2. Results Analysis.** To verify the feasibility of our algorithm, multiple face images were collected from WIDER FACE dataset, i-bug face dataset, and AFW dataset and subjected to experiments using TensorFlow + AlexNet CNN. The experiments were carried out in



FIGURE 12: Rendering of  $Seed_N$ .



FIGURE 15: PPSA results.



FIGURE 13: Rendering of  $Seed_N$  and  $S'$ .



FIGURE 16: Local results at  $\omega \leq \rho$ .



FIGURE 14: LAP results.



FIGURE 17: Local results at  $\omega > \rho$ .

the environment of Intel® Core i9-9900K CPU @ 3.60 GHz, 32G memory, GTX 21080TI GPU, and Windows 10. The privacy budget was set to 1, 2, 3, 4, and 5, in turn. The face recognition performance was measured by precision, recall, and  $F1$ -score. The experimental results are displayed in Figures 18–26.

The experimental results show that the PPSA achieved different results on images from different datasets. The

difference arises from the varied image sizes of different datasets:  $760 * 1000$  in WIDER,  $500 * 800$  in i-bug, and  $1000 * 1500$  in AFW. In addition, the different background complexities between the face datasets also influence the operation effects of LAP, sliding window publication (SWP) algorithm, and region growing publication (RGP). By contrast, the operation of the PPSA is not affected by image size and background complexity.

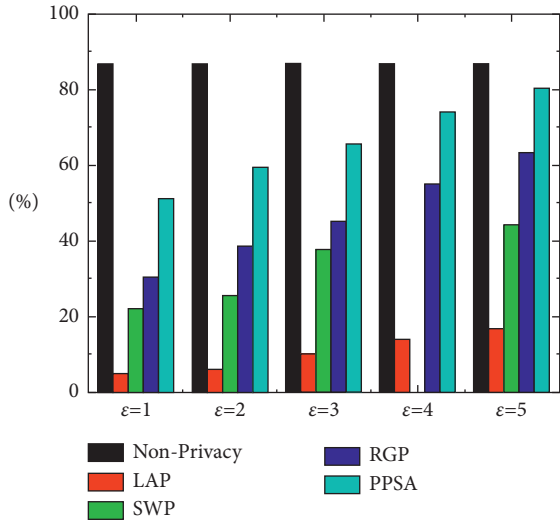


FIGURE 18: WIDER and precision.

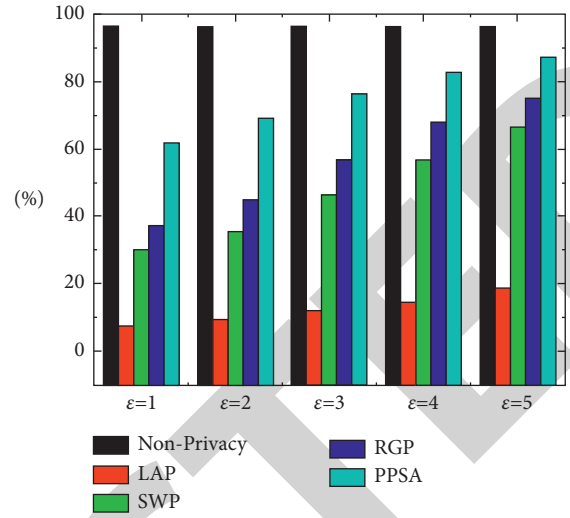


FIGURE 21: I-bug and precision.

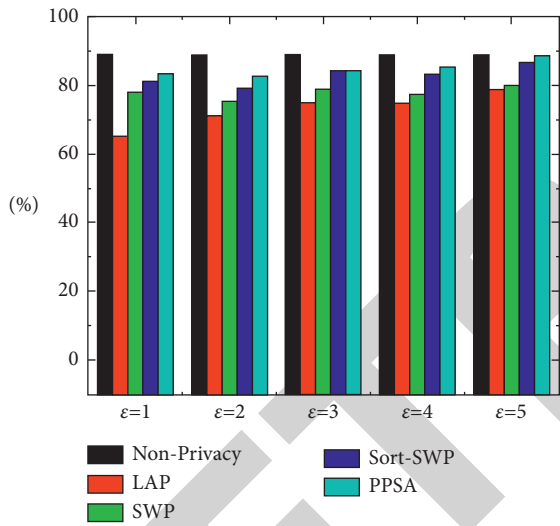


FIGURE 19: WIDER and recall.

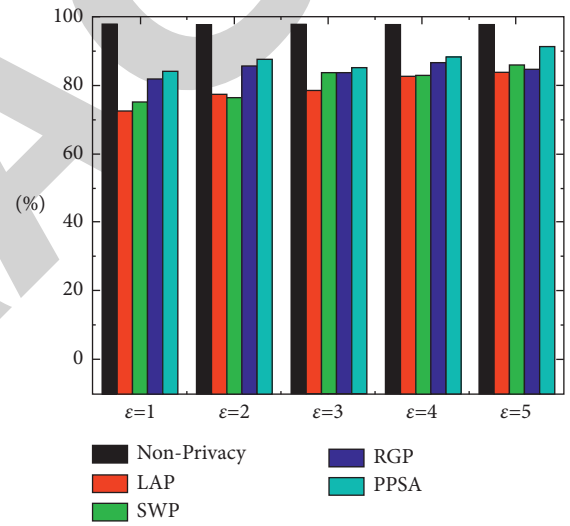


FIGURE 22: I-bug and recall.

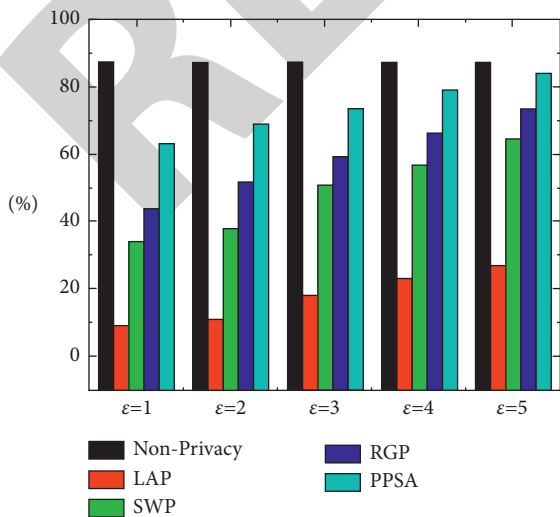


FIGURE 20: IDER and F1-score.

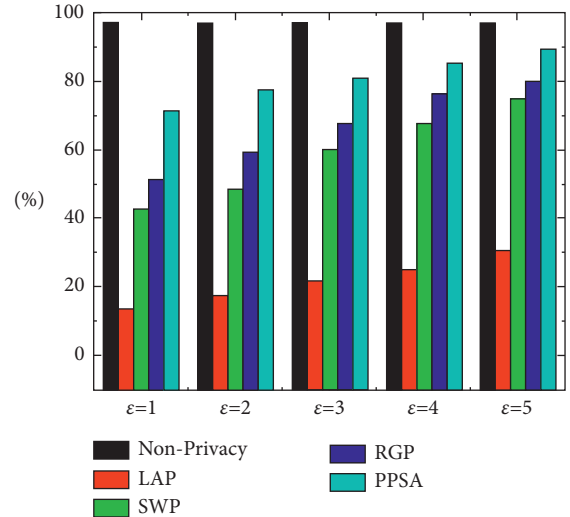


FIGURE 23: I-bug and F1-score.

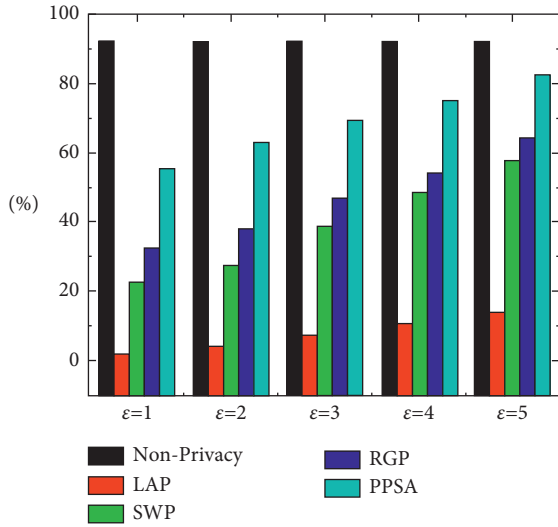


FIGURE 24: AFW and precision.

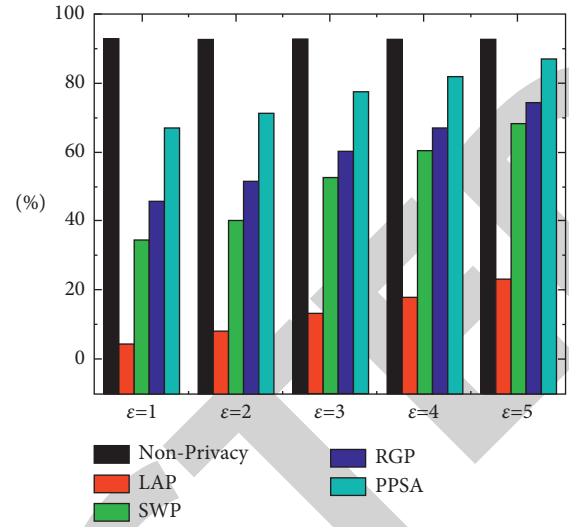


FIGURE 26: AFW and F1-score.

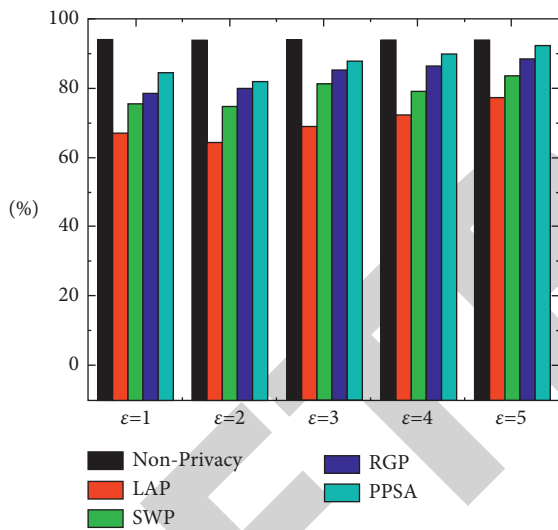


FIGURE 25: AFW and recall.

From the experimental results, it can be learned that the operation results of the PPSA mainly depend on the number of estimated faces  $\rho$  and  $S'$ . In the images from i-bug, there are relatively few faces in each image, but the face area takes up a large portion of the entire image. In the images from WIDER, there are many faces in each image, and the face area takes up a large portion of the entire image. In the images from the AFW, there are relatively few faces in each image, but the face area takes up a small portion of the entire image, for the images tend to be large. Therefore, it is expected that the PPSA's privacy protection effect is positively correlated with  $\omega$ , and negatively with  $S'$ . The prediction is consistent with the experimental results.

## 5. Conclusions

To protect the privacy of multiface images, this paper combines face detection, regional growth, with the Laplace mechanism of differential privacy to add noises to local sensitive areas in face images and realizes the privacy protection for the local sensitive areas in multiface images under an interactive framework. Compared with the LAP, SWP, and RGP, the proposed PPSA can effectively suppress the noise impact on the protected image and improve the availability of the privacy protected image. Moreover, the PPSA applies to images of various sizes, and its error does not increase with the image size.

Although the PPSA can effectively protect the face information in multiface images, the attackers may choose to attack the hair, clothes, and body of persons in the images. Besides, privacy leak may arise from the relationship between persons, and the correlation between each person and the background. The future work will further improve image privacy protection, trying to solve these potential risks.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

The authors thank the project of the National Natural Science Foundation of China (No. 61672179), Natural Science Foundation of Heilongjiang Province (No.



LH2021D022, 2019F004), and Fundamental Research Funds in Heilongjiang Provincial Education Department (No. 145109216).

## References

- [1] B. C. M. Fung, K. Wang, and P. S. Yu, "Anonymizing classification data for privacy preservation," *IEEE Transactions on Knowledge and Data Engineering*, vol. 19, no. 5, pp. 711–725, 2007.
- [2] X. Xiao and Y. Tao, "Anatomy: simple and effective privacy preservation," in *Proceedings of the 32nd IntConf on Very Large Data Bases*, pp. 139–150, ACM, Seoul, Korea, September, 2006.
- [3] T. Li, N. Li, J. Zhang, and I. Molloy, "Slicing: a new approach for privacy preserving data publishing," *IEEE Transactions on Knowledge and Data Engineering*, vol. 24, no. 3, pp. 561–574, 2012.
- [4] M. Terrovitis, N. Mamoulis, J. Liagouris, and S. Skiadopoulos, "Privacy preservation by diSASociation," *Proceedings of the VLDB Endowment*, vol. 5, no. 10, pp. 944–955, 2012.
- [5] L. Sweeney and K. Anonymity, "A model for protecting privacy," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 10, no. 5, pp. 557–570, 2002.
- [6] C. Dwork, "Differential privacy," in *Proceedings Of the 33rd International Colloquium On Automata Languages And Programming*, pp. 1–12, Berlin, 2006.
- [7] C. Dwork, "Differential Privacy: a survey of results," in *Proceedings of the 5th International Conference on Theory and Applications of Models of Computation*, pp. 1–19, Xi'an, China, April, 2008.
- [8] C. Dwork, "The differential privacy Frontier (Extended Abstract)," *Theory of Cryptography*, in *Proceedings of the 6th Theory of Cryptography Conference*, pp. 496–502, San Francisco, CA, USA, March, 2009.
- [9] C. Dwork and J. Lei, "Differential privacy and robust statistics," in *Proceedings of the 41st Annual ACM Symposium on Theory of Computing*, pp. 371–380, Bethesda, MD, USA, January, 2009.
- [10] C. Dwork, "Differential privacy in new settings," in *Proceedings of the Symposium On Discrete Algorithms (SODA)*, January, 2010.
- [11] C. Dwork, "The promise of differential privacy," A Tutorial on Algorithmic Techniques," in *Proceedings of the 52nd Annual IEEE Symposium on Foundations of Computer Science*, Palm Springs, CA, USA, October, 2011.
- [12] C. Dwork, F. McSherry, K. Nissim, and A. Smith, "Calibrating noise to sensitivity in private data analysis," in *Proceedings of the 3rd Theory of Cryptography Conference*, pp. 265–284, New York NY USA, March, 2006.
- [13] F. McSherry and K. Talwar, "Mechanism design via differential privacy," in *Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science*, pp. 94–103, Providence, Rhode Island, USA, October, 2007.
- [14] F. McSherry, "Privacy integrated queries: an extensible platform for privacy-preserving data analysis," *Communications of the ACM*, vol. 53, no. 9, pp. 89–97.
- [15] A. Roth and T. Roughgarden, "Interactive privacy via the median mechanism," in *STOC '10: Proceedings of the forty-second ACM symposium on Theory of computing*, pp. 765–774, New York, NY, USA, June 2010.
- [16] M. Hardt and G. N. Rothblum, "A multiplicative weights mechanism for privacy-preserving data analysis," in *Proceedings of the Foundations Of Computer Science (FOCS), 2010 51st Annual IEEE Symposium*, Las Vegas, NV, USA, October, 2010.
- [17] A. Gupta, A. Roth, and J. Ullman, "Iterative Constructions and Private Data Release," in *Proceedings of the Theory of Cryptography*, pp. 339–356, Berlin Heidelberg, July, 2012.
- [18] L. Fan and L. Xiong, "Real-time Aggregate Monitoring with Differential Privacy," in *Proceedings of the Acm International Conference on Information & Knowledge Management*, Glasgow Scotland, UK, October, 2012.
- [19] G. Kellaris, S. Papadopoulos, X. Xiao, and D. Papadias, "Differentially private event sequences over infinite streams," *Proceedings of the Vldb Endowment*, vol. 7, no. 12, pp. 1155–1166, 2014.
- [20] X. Xiao, G. Wang, and J. Gehrke, "Differential private via wavelet transforms," *IEEE Transactions on Knowledge and Data Engineering*, vol. 23, no. 8, pp. 1200–1214, 2014.
- [21] J. Xu, Z. Zhang, X. Xiao, Y. Yang, G. Yu, and M. Winslett, "Differentially private histogram publication," *The VLDB Journal*, vol. 22, no. 6, pp. 797–822, 2013.
- [22] C. Li, G. Miklau, M. Hay, A. McGregor, and V. Rastogi, "The matrix mechanism: optimizing linear counting queries under differential privacy," *The VLDB Journal*, vol. 24, no. 6, pp. 757–781, 2015.
- [23] C. Li, M. Hay, G. Miklau, and Y. Wang, "A data- and workload-aware algorithm for range queries under differential privacy," *Proceedings of the VLDB Endowment*, vol. 7, no. 5, pp. 341–352, 2014.
- [24] X. J. Zhang, X. J. Zhang, C. C. Fu, and X. F. Meng, "Facial image publication with differential privacy," *Journal of Image and Graphics*, vol. 23, no. 9, pp. 1305–1315, 2018.
- [25] C. Liu, J. Yang, W. Zhao, Y. Zhang, J. Li, and C. Mu, "Face image publication based on differential privacy," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 6680701, 20 pages, 2021.
- [26] C. Liu, J. Yang, W. Zhao et al., "Differential privacy protection of face images based on region growing," *Traitement du Signal*, vol. 38, no. 5, pp. 1385–1401, 2021.
- [27] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [28] Y. Sun, X. Wang, and X. Tang, "Deep convolutional network cascade for facial point detection," in *Proceedings of the Computer Vision and Pattern Recognition*, pp. 3476–3483, Portland, OR, USA, June, 2013.
- [29] Z. Zhang, P. Luo, C. C. Loy, and X. Tang, "Facial landmark detection by deep multitask learning," in *Proceedings of the European Conference on Computer Vision*, pp. 94–108, Pasadena, CA, UK, September, 2014.
- [30] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint face detection and alignment using multitask cascaded convolutional networks," *IEEE Signal Processing Letters*, vol. 23, no. 10, pp. 1499–1503, 2016.
- [31] Y. Wu, T. Hassner, K. Kim, G. Medioni, and P. Natarajan, "Facial landmark detection with tweaked convolutional neural networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 12, pp. 3067–3074, 2017.
- [32] N. Y. Ke and R. Sukthankar, "PCA-SIFT: a more distinctive representation for local image descriptors," in *Proceedings of the IEEE Computer Society Conference on Computer Vision & Pattern Recognition*, Washington, DC, USA, July, 2004.
- [33] C. Liu, J. Yang, and J. Q. Wu, "Web intrusion detection system combined with feature analysis and SVM optimization," *EURASIP Journal on Wireless Communications and Networking*, vol. 2020, no. 1, pp. 1–11, 2020.

## Retraction

# Retracted: A Graph Partition-Based Large-Scale Distribution Network Reconfiguration Method

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] Y. Sha, "A Graph Partition-Based Large-Scale Distribution Network Reconfiguration Method," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3169065, 13 pages, 2022.



## Research Article

# A Graph Partition-Based Large-Scale Distribution Network Reconfiguration Method

Yuanxia Sha 

*Department of Mathematical Sciences, Daqing Normal University, Daqing 163712, Heilongjiang, China*

Correspondence should be addressed to Yuanxia Sha; shayuanxia@dqnu.edu.cn

Received 17 January 2022; Accepted 1 February 2022; Published 13 March 2022

Academic Editor: Daqing Gong

Copyright © 2022 Yuanxia Sha. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article focuses on the analysis of large-scale distribution network reconstruction fused with graph theory and graph partitioning algorithms. Graph theory and graph segmentation algorithms have been rushed by many researchers in the fields of medicine, drone, and neural network. It is a newcomer in the field of computer vision, which can not only realize the division in color but also divide it by image data. The distribution network is also indispensable for new energy, electric machines, but the traditional distribution network has many problems, such as not suitable for distributed power access and excessive network loss. To improve the performance of distribution networks and reduce network losses, this paper A multi-division model for distribution network construction and reconstruction is established, and a graph theory-based division algorithm method is proposed to effectively solve the problem of feeder-to-feeder reconstruction during large-scale distribution in distribution networks. Through its superconductivity phenomenon and the characteristics of clustering algorithm division, this paper uses formulas to show its division principle and gives examples of various distribution network reconstruction algorithms to explore which method of improvement can improve the performance of the distribution network and reduce network losses. The number of iterations is also strictly considered, and the value is taken after multiple iterations to reduce the error. Through the distribution network calculation example, the network loss reduction value is obtained, and the distribution network fault repair model is exemplified. The picture is used to briefly describe the process of distribution network reconstruction and find that the faults of the distribution network can be quickly located and isolated through the FTU, and quickly repaired. Finally, in order to reduce the network loss, reduce the load of power flow calculation, and solve the problem of local optimization, a JA-BE-JA optimization algorithm based on large-scale distribution network reconfiguration is proposed. The mixed sampling method is preferred to test the number of divisions in the four states, and the parameters are selected to test the performance of the improved annealing simulation algorithm, and the conclusion is drawn as follows: the improved graph segmentation algorithm has strong robustness, can avoid local optimization of graph data, and can reduce network loss. Compared with traditional distribution network reconstruction methods, the network loss can be reduced to 454.3 KW, which can be optimized by 10.68% compared with the initial network loss.

## 1. Introduction

As a rigorous and scientific theory, graph theory has now been widely used in various fields. The first successful application of graph theory occurred in the 1990s. Some scholars successfully applied it in the research of image segmentation, which attracted the attention and discussion of many scholars at that time. In the development of modern society, the theory of graph theory has also been applied more effectively and broadly, such as texture synthesis, image restoration, and the complicated construction and

reconstruction of power distribution networks in modern society. The emergence of graph theory has become extremely significant, which effectively helps in large-scale problems such as distribution networks. Therefore, by establishing a multipartition model for the construction and reconstruction of the distribution network, this paper proposes a partition algorithm based on graph theory to effectively solve the problem of reconfiguration between feeders during large-scale distribution in the distribution network. The proposal of this algorithm greatly reduces the computational pressure in network reconstruction and can

effectively arrive at the most reasonable solution based on the global distribution network. Since graph theory itself is a branch subject based on mathematics, the algorithm proposed in the article is a new and effective graph theory algorithm based on rigorous mathematical proofs. In the reconstruction of the distribution network, the image segmentation technology plays an extremely important role. The topology of the network is changed by constantly switching the switch states of the tie switch and the section switch, so as to regulate the flow of power in the entire distribution network. The branch exchange method, optimal flow mode method, genetic algorithm, etc., are all effectively applied in the construction of the distribution network. The improved graph segmentation algorithm can effectively reduce the network loss caused by the operation of the distribution network. In the subsequent reconstruction of the distribution network, graph segmentation algorithms have great research significance.

The purpose of this article is to study the analysis of large-scale distribution network reconstruction fused with graph theory and graph partitioning algorithms and to propose optimization algorithms to reconstruct the distribution network. The algorithms are tested before and after optimization through experiments and applied in the environment of distribution network reconfiguration, collecting the number of edge cuts in the distributed segmentation of 4 image datasets with data sizes ranging from small to large by the JA-BE-JA algorithm based on the BSP-Spark platform. Then, the edge cut numbers of the JA-BE-JA algorithm are collected based on the BSP-PeerSim platform in the same dataset, and they are all collected after 1200 supersteps to ensure the accuracy of the experiment, and the strong robustness of this algorithm and the characteristics of avoiding local optimization are obtained. A 16-bus distribution network is selected for testing, which contains 3 feeder parts, indicating that this algorithm can be applied to large-scale distribution networks and can reduce network losses. At the same time, it also shows that, in dealing with the problem of large-scale distribution network reconfiguration, the optimization algorithm proposed in this paper has a better effect, has a certain guiding significance, and has both theoretical and practical significance.

The improved algorithm designed in this article has a good application prospect in large-scale distribution network reconstruction. Based on the image segmentation method of graph theory-full variational approximation of the Ginzburg-Landau functional of graph clustering, the power flow algorithm of the distribution network reconstruction and the distribution network fault recovery model to perform certain segmentation of the distribution network and the graph theory explanation, the two do have certain commonalities. Then through experiments, the improved algorithm of graph theory and graph segmentation is integrated into the distribution network reconstruction. In order to avoid errors, there are some details in the experiment, such as the choice of encoder and the choice of iteration times. Finally, it is concluded that this graph partitioning improved algorithm can be better applied to a large-scale distribution network environment, and the

network loss and robustness are all optimized, which can play a certain guiding role in future development, and it can also be used for reference in other fields.

## 2. Related Work

The application of graph theory and graph segmentation algorithms in various fields has extremely optimistic development prospects, and all fields have never stopped the research and exploration of graph theory and graph segmentation algorithms in order to have a brighter development prospect. Li and Yang paid attention to the adaptive pinning synchronization problem of random complex dynamic network (CDN); based on the algebraic graph theory and Lyapunov theory, they deduced the design conditions of the pinning controller and performed a strict convergence analysis on the synchronization error of probabilistic meaning. Compared with the existing results, due to the use of graph theory, the topology of the random CDN remains unknown [1]. Keown et al. used a selective high-quality data subset from autism brain imaging data exchange (including 111 ASD and 174 typical developmental participants) and several graph theory indicators. They preprocessed and analyzed resting state functional magnetic resonance imaging data to detect low-frequency inherent signal correlation and concluded that compared with typical development participants, the Rand index (reflecting the similarity between the network organization and the normative network set) of ASD participants was significantly lower [2]. Glaria et al. stated that there are many benefits of compressing real-world graphs, such as improving or enabling visualization in small memory devices, graph query processing, community search, and mining algorithms. This work proposes a novel compact representation for real sparse and clustered undirected graphs. This method uses a fast algorithm to list all the largest cliques and defines a clique graph according to its largest clique [3]. Li et al. showed that the existing vector-based methods of machine learning usually use vector-based features to represent the program, but it is easy to ignore the control information between the basic block and the partition on the path other than the critical path. And they proposed a novel graph-based thread partitioning method to overcome these two bottlenecks. They used graphs to characterize the program, integrated features and control information, and successfully extracted a good partition scheme [4]. Ghasemi said that the problem of power distribution system reconfiguration is a complex optimization process to find a structure with the least loss, which needs to meet the satisfaction of both consumers and the power distribution system company. One of the most important parameters in this regard is to improve the reliability of the system. On the one hand, this parameter improves the satisfaction of electricity consumption, and on the other hand, it improves the economic benefits of power distribution enterprises [5]. Hong H et al. proposed a distribution network reconfiguration method based on a directed graph considering distributed generation. Two reconfiguration scenarios are considered: operating mode adjustment with the goal of minimizing active

power loss (case I) and service restoration with the goal of restoring the load to the greatest extent (case II). These two situations are modeled as a mixed integer quadratic programming problem and a mixed integer linear programming problem [6]. Although these studies have reached certain guiding conclusions, there are unavoidable errors, or insufficient optimization effects, insufficient demonstrations, etc., which need to be further improved.

### 3. Distribution Network Reconstruction Model and Graph Theory-Graph Segmentation Method

*3.1. Image Segmentation Algorithm Based on Graph Theory-Full Variational Approximation of the Ginzburg-Landau Functional of Graph Clustering.* The Ginzburg-Landau function is based on the phase transition theory, which combines superconducting phenomena such as quantum mechanics, electrodynamics, and quantum mechanics of superconductors to obtain related nonlinear partial differential equations [7]. Formula (1) expresses the relationship between phase change and phase separation. Letting  $S$  represent the open subset of  $R^t$ , the parameter  $u > 0$  is the spatial scale:

$$GL(v) := \frac{u}{2} \int_S |\nabla v(a)|^2 dx + \frac{1}{u} \int_S W(v(a)) dx. \quad (1)$$

The form of the Ginzburg-Landau function is a state transition equation, is closely related to the state transition equations such as Allen-Cahn and Chafee-Infante, usually because of the connection between the Ginzburg-Landau function and the total variational seminorm, and is regarded as the second approximation of the total variation, so it can solve the nonsmooth total variation minimization question [8]. The formula used when processing the image is

$$B(v) = GL(v) + \lambda f(v, v_0), \quad (2)$$

where  $f(v, v_0)$  is the distance function of the reconstructed image  $v$  and the given image  $v_0$ . The statistics of the model are closely related to the physical assumptions. Among them, the parameter  $\lambda > 0$  can control the influence of data fitting on regularization. Applying the  $L^2$  gradient descent method, the partial differential equation at this time can be known:

$$f(v, v_0) = \|v - v_0\|_{L^d(S)}^d, \quad d = 1, 2, \quad (3)$$

$$vt = -\frac{\beta GL}{\beta v} = u\Delta v - \frac{1}{u}W'(v) - \lambda \frac{\beta f}{\beta v}.$$

Because of the existence of the distance function  $f(v, v_0)$ , here we add a forced term  $\lambda\beta f/\beta v$ . Since  $B(v)$  is a nonconvex function, there cannot be a unique solution at this time, and the result depends on the initial conditions. The prior information of the image can be used as the initial input of the segmentation algorithm to achieve the goal of more accurate image segmentation in graph theory. Then, the segmented image can be classified with high-dimensional data, and then  $I := \{a = (a_1, a_2) \in C^2: 0 \leq a_1 \leq N-1; 0 \leq a_2 \leq M-1\}$  is set

to represent a rectangular image with  $K = N \times M$  pixels. The image neighborhood of  $a$  can be represented as set  $M(a), a \in I$ :

$$M(a) := \{b \in I: |a_1 - b_1| \leq \varepsilon \text{ and } |a_2 - b_2| \leq \varepsilon\}. \quad (4)$$

When  $\varepsilon \in M$ , there are  $(2\varepsilon + 1) \times (2\varepsilon + 1)$  size pixels centered on  $a$  in  $M(a)$ ; for  $K \in M$ , we can associate the feature vector  $c \in R^K$  selected by the  $a$  and  $M(a)$  neighborhoods. Let  $X$  be the pixel node,  $B$  the relationship between the vertices mapped to the edge of the graph,  $\omega$  the weight, and  $W$  the similarity matrix, then

$$\begin{aligned} \omega(a_z, a_i) &= \omega(a_i, a_z), \omega(a_z, a_i) > 0, \\ \omega(a_z, a_i) &= \hat{\omega}(\gamma(a_z), \gamma(a_i)) = \hat{\omega}(c_z, c_i), \end{aligned} \quad (5)$$

where  $\gamma$  is the characteristic function. At this time, in order to make the symbols consistent, let  $\omega(a_z, a_i) = \hat{\omega}(a_z, a_i)$ , facing the segmentation problem, take the weight of the edge connected to node  $x_z$  and the degree  $d_z$  of the vertex to obtain a diagonal matrix  $D$  of  $m \times m$ :

$$d_z = \sum_{i=1}^m \omega_{zi},$$

$$D = \begin{bmatrix} d_1 & \cdots & \cdots & \cdots \\ \cdots & d_2 & \cdots & \cdots \\ \vdots & \vdots & \ddots & \vdots \\ \cdots & \cdots & \cdots & d_m \end{bmatrix}. \quad (6)$$

When  $X$  is used to represent the space of  $X \rightarrow R$  of all functions,  $X \rightarrow X$  is its Laplacian  $L$ :

$$Lv(a) = \sum_{b \in X} \omega(a, b)(v(a) - v(b)), \quad a \in X,$$

$$L(a, b) := \begin{cases} d(a), & \text{if } (a = b) \\ 0 - \omega(a, b), & \text{otherwise} \end{cases}. \quad (7)$$

The related quadratic form  $H$  of  $L$  can be expressed as

$$H(v, Lv): \frac{1}{2} \sum_{a, b \in X} \omega(a, b)(v(a) - v(b))^2. \quad (8)$$

The non-negative real eigenvalue of  $L$  is  $\{\lambda_z\}_{z=1}^a$ , of which  $0 = \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_p$ , and under the large-scale limit,  $L$  can be appropriately scaled to achieve high stability of convergence. At this time, Tulaplas can be normalized:

$$L_p := D^{-1/2} L D^{-1/2} = I - D^{-1/2} W D^{-1/2}. \quad (9)$$

At this time, the matrix  $L_p$  presents a symmetric state, redefining the discrete Ginzburg-Landau function:

$$GL_d(v) := \frac{u}{2} H(v, L_p v) + \frac{1}{u} \sum_{a \in X} W(v(a)) + \sum_{a \in X} \frac{\chi(a)}{2} (v(a) - v_0(a))^2, \quad (10)$$

$$\omega(a, b) = \exp(-\|\gamma(a) - \gamma(b)\|^2 / \sigma^2), \quad \sigma > 0.$$

Because this image is an undirected graph,  $\omega(a, b)$  will be in a symmetrical state, that is,  $\omega(a, b) = \omega(b, a)$ . Here, in order to reduce the computational complexity, some feature vectors will be used to construct vectors for graph clustering, and then clustering algorithms will be used to divide their feature vectors to segment the image [9]. The feature vector value is related to the gray or RGB (red, green, blue) intensity value and the texture feature of the neighborhood.

**3.2. Power Flow Algorithm for Distribution Network Reconstruction.** When the distribution network is being reconfigured, when its topological structure is radial, the power flow algorithm can be used to calculate the operation of the distribution network, network losses, and voltage offsets [10]. At present, the power flow algorithms suitable for distribution networks are mainly divided into the Newton method, branch method, and bus method. However, there are certain limitations, and the power flow algorithm is not applicable to the distribution network in some cases. For example, the distribution network and the transmission network have different line parameters, different resistance and reactance ratios, or different numbers of OV nodes. The distribution network will have problems such as frequent line changes and complex structure.

Therefore, there will be more power flow algorithms used in the transmission grid than in the distribution network. Based on the increasing attention of the distribution network, improving the convergence, timeliness, and simplicity of the power flow algorithm has become a major goal of the research [11]. Figure 1 is a simple radial distribution network circuit diagram, and its parameters are explained as follows.

The radial distribution network mainly relies on the current quick-break protection at the line outlet, circuit breaker, load switch, and recloser to protect the circuit.

**3.2.1. Improving the Basic Principle of the Cow Pull Method.** Because it has a better solution effect for nonlinear equations, it is often used in power flow calculations of distribution networks. The Newton-Raphson method voltage polar coordinate formula is

$$\dot{U}_k = U_k < \beta_k = U_k (\cos \beta_k + i \sin \beta_k). \quad (11)$$

Let the distribution network have  $x$  OL nodes, denoted by  $1, 2, 3, \dots, x$ , respectively, and  $y - x - 1$  PV nodes, denoted by  $x + 1, x + 2, \dots, y - 1$ , plus a balance node, then all nodes satisfy

$$\begin{cases} \Delta O_k = O_{kr} - O_k = O_{kr} - V_k \sum_{i=1}^y V_i (W_{ki} \cos \beta_{ki} + M_{ki} \sin \beta_{ki}) = 0; k = 1, 2, \dots, y - 1, \\ \Delta L_k = L_{kr} - L_k = L_{kr} - V_k \sum_{i=1}^y V_i (W_{ki} \sin \beta_{ki} - M_{ki} \cos \beta_{ki}) = 0; k = 1, 2, \dots, x. \end{cases} \quad (12)$$

According to Taylor's expansion, the modified formula of Newton-Raphson's method can be obtained as

$$\begin{bmatrix} \Delta O \\ \Delta L \end{bmatrix} = - \begin{bmatrix} S & D \\ F & G \end{bmatrix} \begin{bmatrix} \Delta \beta \\ U_{H_2}^{-1} \Delta U \end{bmatrix}. \quad (13)$$

The small distance between the actual nodes will only form a small voltage difference between adjacent nodes, so in the calculation, the ground admittance parameter can be ignored. At this time, if  $k \neq i$ , then

$$\begin{aligned} S_{ki} &\approx \frac{\beta O_k}{\beta \theta_i} = U_k U_i M_{ki} \cos \theta_{ki}, \\ D_{ki} &\approx U_i \frac{\beta O_k}{\beta U_i} = -U_k U_i W_{ki} \cos \theta_{ki}, \\ F_{ki} &\approx U_i \frac{\beta L_k}{\beta \theta_i} = U_k U_i W_{ki} \cos \theta_{ki}, \\ G_{ki} &\approx U_i \frac{\beta L_k}{\beta \theta_i} = U_k U_i M_{ki} \cos \theta_{ki}. \end{aligned} \quad (14)$$

If  $k = i$ , then

$$\begin{aligned} S_{ki} &\approx \frac{\beta O_k}{\beta \theta_i} = -U_k \sum_{\substack{i=1 \\ i \neq k}}^y U_i M_{ki} \cos \theta_{ki}, \\ D_{ki} &\approx U_k \frac{\beta O_k}{\beta U_k} = U_k \sum_{\substack{i=1 \\ i \neq k}}^y U_i W_{ki} \cos \theta_{ki}, \\ F_{ki} &\approx U_k \frac{\beta L_k}{\beta \theta_k} = U_k \sum_{\substack{i=1 \\ i \neq k}}^y U_i W_{ki} \cos \theta_{ki}, \\ G_{ki} &\approx U_k \frac{\beta L_k}{\beta \theta_k} = U_k \sum_{\substack{i=1 \\ i \neq k}}^y U_i W_{ki} \cos \theta_{ki}. \end{aligned} \quad (15)$$

The iterative steps of improving the Niu La method are as follows: First, the voltage of the balance node and its phase angle are set, and the parameters of the other node lines are initialized. Then, the calculated parameter values are

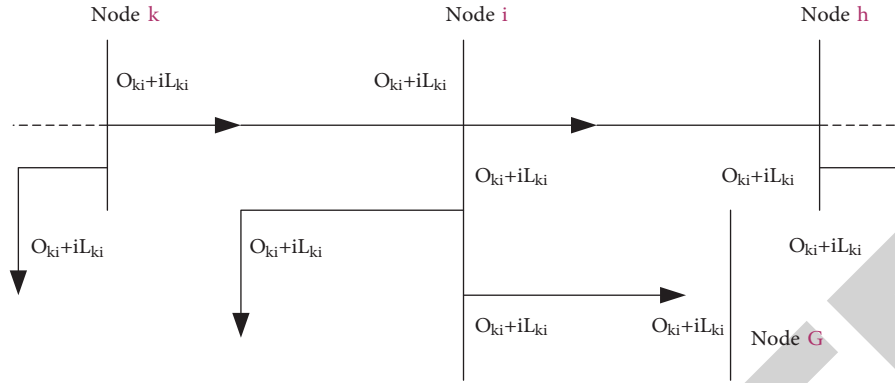


FIGURE 1: Circuit diagram of a typical radial distribution network.

calculated for active and reactive power corrections, and the accuracy of the power flow calculation is judged until it meets the calculation requirements. Then, the above parameters are substituted into the Jacobian matrix to calculate the voltage amplitude correction amount and phase angle correction amount [12]. Let the number of iterations at this time be  $h$ , then

$$U_k^{(h+1)} = U_k^{(h)} + \Delta U_k^{(h)}; \theta_k^{(h+1)} = \theta_k^{(h)} + \Delta \theta_k^{(h)}. \quad (16)$$

Then, the above work is gradually repeated, and the iteration continues.

**3.2.2. Distribution Network Reconstruction with DG.** The reconstruction step is to initialize various parameters first and number the goals of the distribution network reconstruction and the simplified branches, reinitialize part of the variable value and judge the shape, determine whether it is radial and decide whether to reapply the value, then find the variable of another part of the branch of the distribution network that has been disconnected, and iterate the speed and position. After that, the hierarchical matrix is constructed and combined with the distributed power processing process, and it is pushed back to the power flow calculation before it is completed. Then, the next step is to compare the objective function values to determine the optimal population particle value, continue to iterate and judge the distribution network structure until it is radial and reaches the number of iterations, and finally output the optimal population particle value [13]. The steps are shown in Figure 2.

Here, we will take the IEEE69-node system as an example of distribution network reconstruction. At this time, there are 69 nodes, numbered from 1 to 69. The solid line is the segment switch, and the dashed line is the tie switch. The total system load in this distribution network topology is  $3814.6 + j2702.2$  kVA, and there are 98 iterations of the algorithm. The forward and backward power flow calculations are used to iteratively converge to achieve the minimum network loss. The topology of the distribution network is shown in Figure 3.

In the case of the reconfiguration of the distribution network without distributed power, the initial topology of

IEEE69 is that the dashed part denotes the contact switch is off and the solid line denotes the section switch is closed [14]. The reconstruction of the IEEE69-node system without DG is shown in Table 1.

At this time, the network loss of the common algorithm is reduced by 52.98% compared with that before reconstruction, while the algorithm in this paper is reduced by 57.08% compared with that before reconstruction, and the performance is increased by 4.1%, and the effect is better. The network loss of the reconfiguration of the distribution network with distributed power sources is shown in Table 2.

It can be seen that reasonable arrangement of grid-connected locations can reduce network loss. At this time, it has been reduced by 17.27% before reconstruction. After reconstruction, it is reduced by 69.74%. It can be seen that, with the DG power grid, the advantage of its topology is obviously higher than that of power grid without DG. That is, DG reconstruction can improve the voltage quality of the system nodes and improve the operation stability of the distribution network [15].

**3.3. Distribution Network Fault Recovery Reconstruction Model.** Normally, the realization principle of the distribution network fault recovery reconstruction is to use the application layer and the decision-making layer of the distribution network. The control system writes the reconstruction algorithm into the distribution network dispatching system, and the relevant FTU in the distribution network transmits real-time information at the section switch and contact switch. Incorporating data into the SCADA system, the faulty area is located, and here the SCADA system is the control center of the distribution network. Commands are sent to the FTU to isolate the faulty area by operating the switch, the self-healing system is used to generate the structure, optimize algorithm calculations, and control the data, and then the fault area recovery process is performed [16]. In this process, the optimization algorithm is the core part of the distribution network reconstruction, and the voltage, current, power, resistance, frequency, and other states of the distribution network need to be set, and give the range of functions in different states to explore the operation state of the distribution network at this

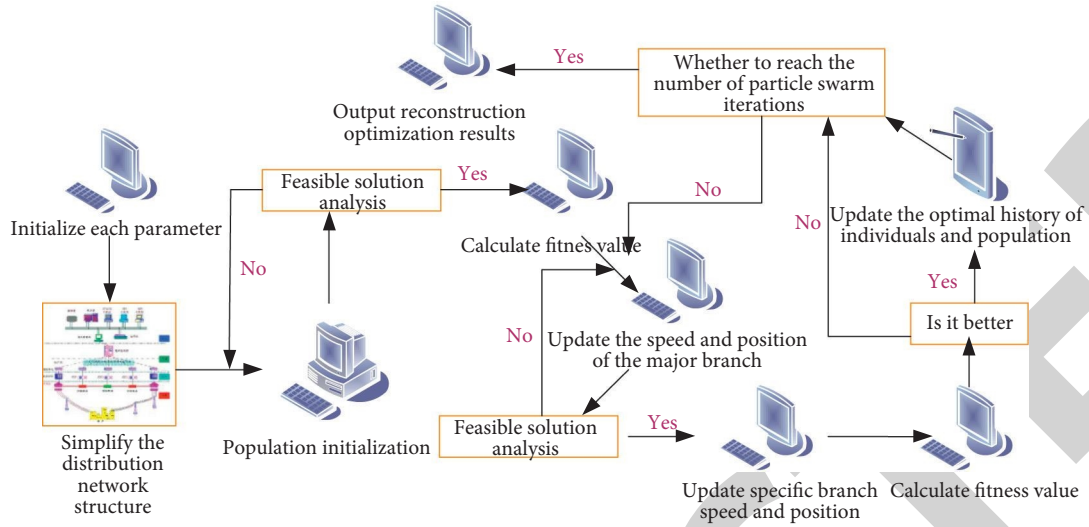


FIGURE 2: Distribution network reconstruction algorithm flow.

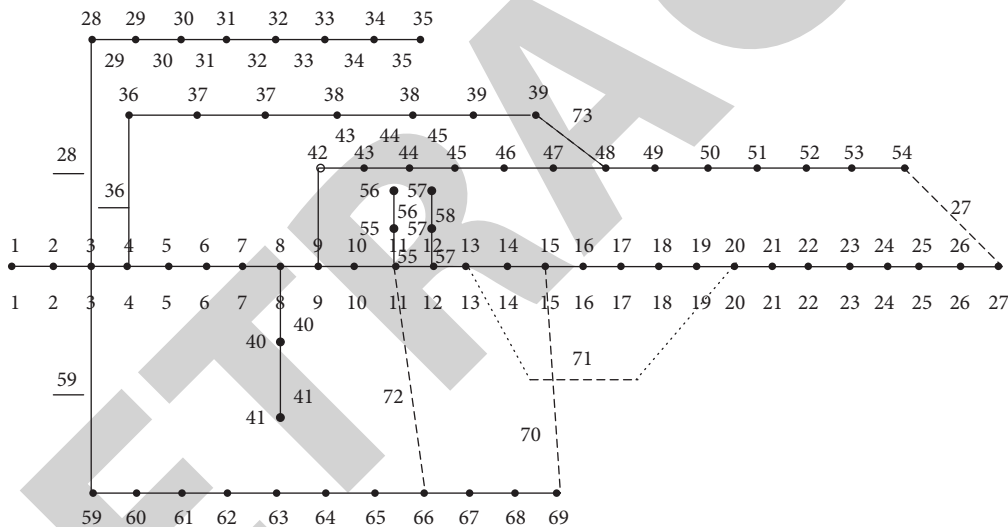


FIGURE 3: IEEE69 distribution network node topology.

TABLE 1: Distribution network reconstruction results without DG grid connection.

State	Disconnect switch position	Lowest node voltage	Network loss (kW)
Before reconstruction	27/70/71/72/73	0.912	231.25
Ordinary algorithm	13/43/50/71/72	0.951	108.74
The algorithm of this paper	14/44/50/71/72	0.946	99.26

TABLE 2: Distribution network reconstruction results when DG is connected to the grid.

State	Disconnect switch position	Lowest node voltage	Network loss (kW)
Before reconstruction	27/70/71/72/73	0.912	231.25
Before reconstruction (including DG)	27/70/71/72/73	0.931	191.32
After reconstruction (including DG)	13/46/51/71/72	0.962	69.98



time, and then carry out the optimization search operation [17]. These states include the power grid emergency state, abnormal state, and realm state. The control process of the distribution network is shown in Figure 4.

In the restoration and reconstruction of the distribution network failure, the FTU can quickly locate and isolate the distribution network failure and quickly repair it.

*3.4. Improved Graph Segmentation Based on Graph Theory BSP-Spark System and Improved JA-BE-JA Algorithm Based on BSP-Spark Cluster.* This experiment will be carried out in the environment of a BSP-Spark distributed cluster composed of 4 PCs. Among the PCs, one is used for the main control system with 8 GB of memory, and three are used as slave computing nodes with 4 GB of memory. The software used for compilation is IntelliJ IDEA, and the program used for language writing is Scala. In order to minimize the experimental errors caused by various factors, the machines here will perform tasks one-to-one, and Spark on Yarn will complete resource management and scheduling in the experiment. The storage system is a distributed file system HDFS [18]. It is known that the BSP-Spark system architecture is an M/S architecture, with a complete master-slave relationship, implemented by 1 master node and 3 child nodes. The node deployment diagram is shown in Figure 5.

If it wants to achieve high-performance distributed partitioning, it must optimize the parameters of the BSP-Spark system, such as memory, parallelism, and data skew. Although Spark's calculation is based on memory, there are also algorithms that do not rely on memory. At this time, a caching mechanism needs to be added. In general, the caching method used is `spark.executor.memory`, and large-scale iterative algorithms need to use a storage mechanism to control the consumption of data access. This article will reduce the set cache to reduce memory consumption, and `System.setProperty` can just solve this problem [19]. Execution efficiency is closely related to Spark parallelism. Too low parallelism will lead to waste of resources, while too high will reduce efficiency. At this time, `spark.default.parallelism` is used to optimize the parallelism to ensure the number of tasks. One of the major problems of distributed systems is data skew. During Spark application development, memory overflow exceptions are usually caused. A relatively large amount of graph data corresponds to a relatively small number of node calculations, which will also cause its execution speed to tilt. At this time, shuffle is used to reduce the number of partitions. When small-scale data appear, broadcast variables can be used, and the time interval can be set by `spark.speculation.interval` [20].

The JA-BE-JA algorithm is a graph partitioning algorithm that combines point partitioning and edge partitioning for load balancing. It comes from the PeerSim environment. Based on the JA-BE-JA algorithm of the BSP-Spark platform, image nodes can be processed asynchronously on a regular basis, and a node only interacts with a small random range of adjacent nodes. The information dissemination between nodes is only through the edges of the image and is based on memory sharing. Its vertices

correspond to a single node host and process threads independently. The image dataset input by the improved JA-BE-JA algorithm is only in  $G(V,E)$  format, and the image data need to be preprocessed before dividing. Preprocessing includes split-transform-storage, and graph data cleaning uses transformation and action operators to generate new edge RDDs and vertex RDDs from the data source. The algorithm flow is shown in Figure 6.

The edge segmentation and point segmentation of the JA-BE-JA algorithm are realized by the `Graph.partitionBY` operator, and the mixed sampling method is preferred for the sampling strategy.

## 4. Results and Analysis

*4.1. Experimental Results and Analysis.* Here, we will collect the edge cuts of the JA-BE-JA algorithm based on the BSP-Spark platform in the distributed segmentation of 4 image datasets with data sizes ranging from small to large, then collect the edge cuts of the JA-BE-JA algorithm based on the BSP-PeerSim platform in the same dataset, and collect them after 1200 supersteps to ensure the accuracy of the experiment [21–23]. The experimental results are shown in Figures 7 and 8.

It can be seen that the size of the vertices and edges of the 4elt graph is small, and the difference in the number of edge cuts is small, while the size of the vibrobox graph is larger, and the trend of the number of edge cuts gradually tends to be flat.

Next, the improved annealing simulation algorithm is analyzed. It is known that the improved BSP-Spark algorithm has better load in large-scale graph segmentation, but it is easy to form a local optimal state. Therefore, this article will propose an improved annealing algorithm based on BSP-Spark and perform performance testing, taking two parameters  $u$  and  $v$  for performance measurement; let  $u$  be the parameter of the edge cut problem of the improved JA-BE-JA algorithm and be the system energy function value. The relationship between  $u$  and JA-BE-JA algorithm in the original BSP-PeerSim and BSP-Spark environments is shown in Figure 9.

It can be seen that the improved annealing simulation algorithm has a smaller edge cut when  $u = 2$ . This is because when this parameter is greater than 1, graph partitioning can reduce the number of communication edges without damaging the original topological structure of the graph and reduce information propagation and iteration time. When the parameter  $u = 3$ , the consumption of information dissemination will increase due to the reduction of the total effect value of the graph. This situation will become more serious as the parameter  $u$  increases. The improved annealing model algorithm can reduce the consumption when  $u = 2$  and reduce the probability of the local optimal problem of the JA-BE-JA algorithm. Let  $v$  be the cooling coefficient, which can represent the cooling rate, and the relationship between it and the edge cut number of the JA-BE-JA algorithm is shown in Figure 10.

It can be seen from the Figure 10 that based on the original BSP-PeerSim environment, the number of edge cuts increases with the increase of the parameter  $v$ , and  $v$  can

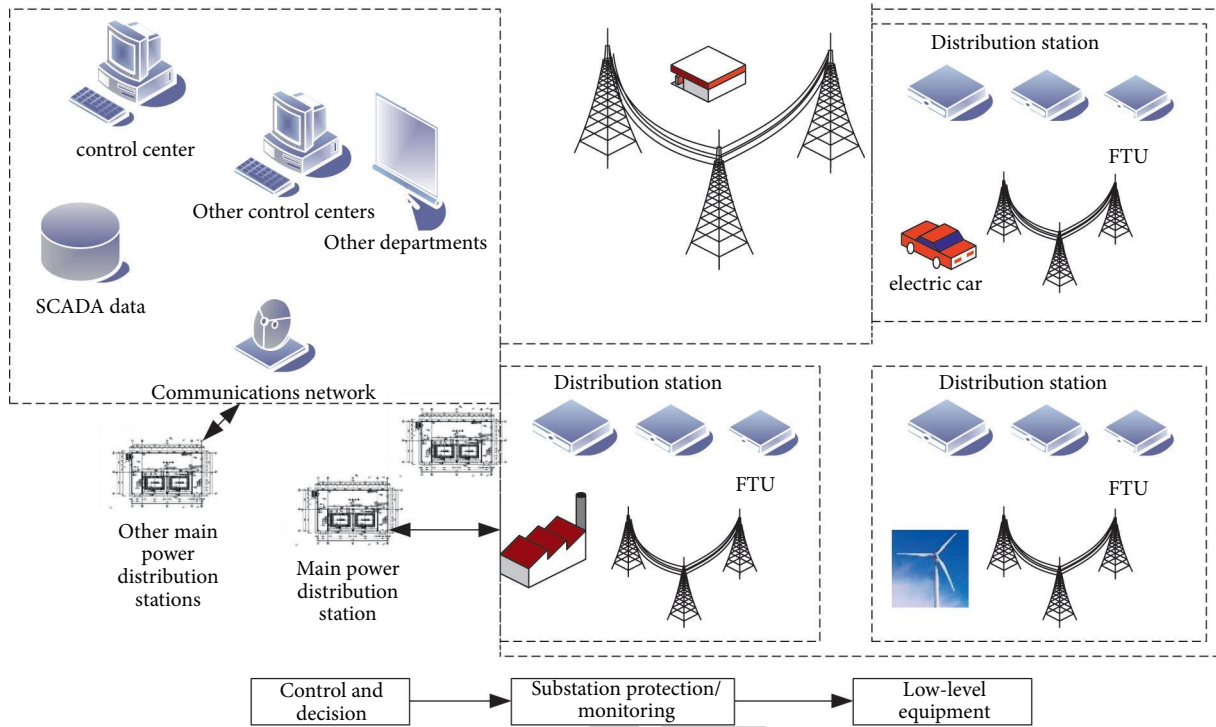


FIGURE 4: Distribution network control process.

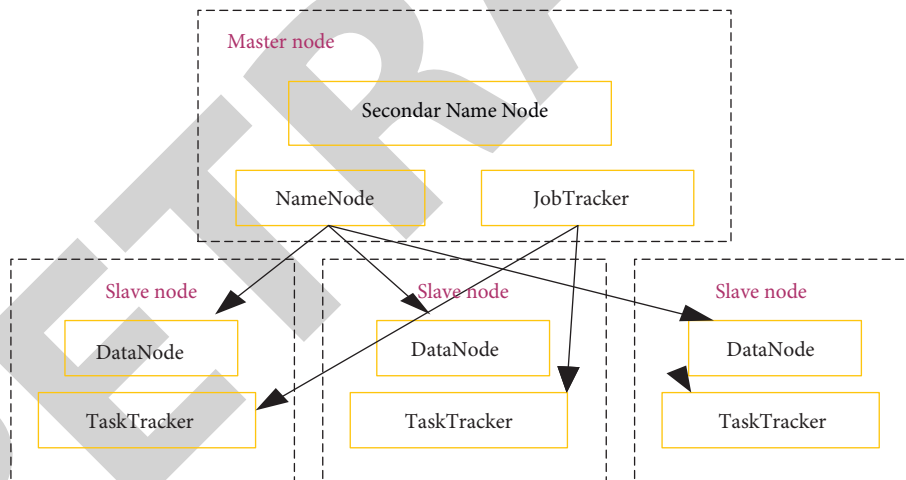


FIGURE 5: Node deployment.

represent the equalization effect of the number of exchanges and the quality of graph division. That is to say, the greater the number of exchanges, the longer the convergence time, and the information dissemination consumption will increase. Based on the BSP-Spark environment, the edge cut value of the obvious graph has robust stability for the parameter  $\nu$ .

In the improved data migration algorithm, local optimization problems can be avoided, and random functions can be used to reduce the number of communication edges. At this time, using different datasets for testing, it can be seen that a 4-way partition will cause 76.3% of the vertex data to migrate, and only 23.7% of the data moved to different partitions.

4.2. Improved Graph Theory and Graph Segmentation Method Cited into the Distribution Network Calculation Examples and Result Analysis. Here, we will use the improved JA-BE-JA algorithm to calculate the distribution network example. The 16-busbar system of the distribution network has 3 tie switches and 3 feeders, as shown in Figure 11.

The tie switch is closed, the weighting operation is performed on the 16-bus distribution system, the eigenvector is solved by the spectral division algorithm, and the critical edge of the above-mentioned distribution system and the node of division 1 are removed [24]. Figure 12 is available.

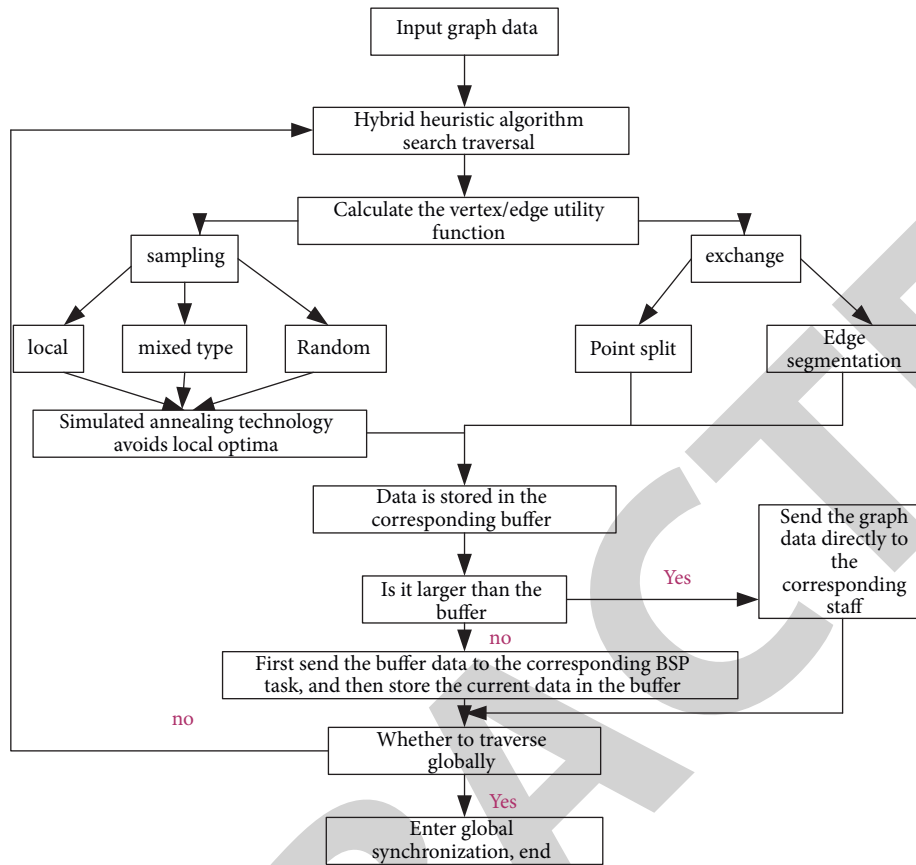


FIGURE 6: Improved JA-BE-JA algorithm flow.

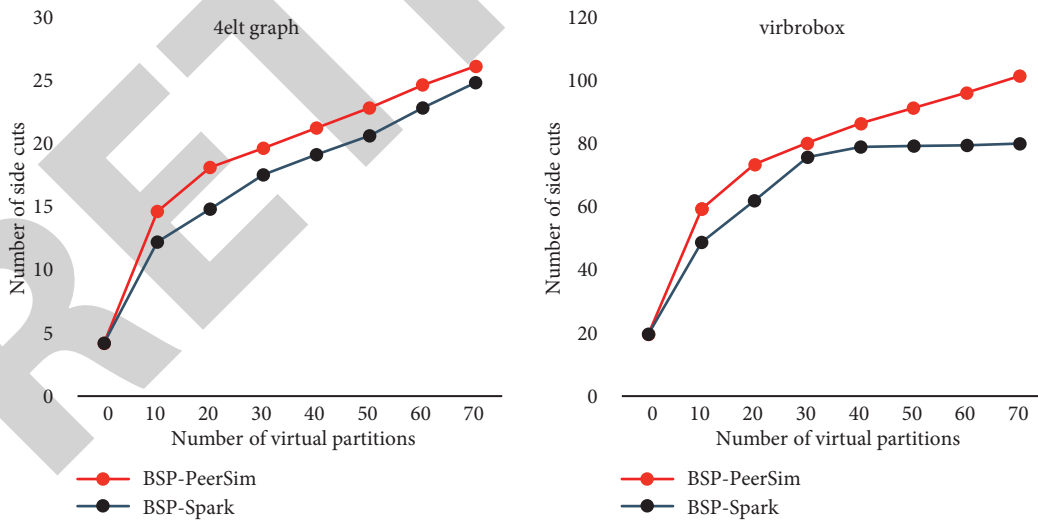


FIGURE 7: Relationship between the two partitions of 4elt graph and vibrobox and the edge cut number of the JA-BE-JA algorithm.

The above figure is a subgraph based on the feature evaluation value after the first scoring and sorting, and its feature evaluation value is shown in Figure 13 at the same time as the feature evaluation value after the second scoring and ranking. The distribution network nodes are divided into groups: group 1

is {1, 2, 4, 5, 6, 7, 8, 9, 11, 12}, group 2 is {3, 10, 13, 14, 15, 16}, which is recorded as division 1, and the rest of the system needs to be divided into two using the improved JA-BE-JA algorithm, that is, division 2: {1, 4, 5, 6, 7, 11} and division 3: {2, 8, 9, 12}, and Figure 13 can be obtained.

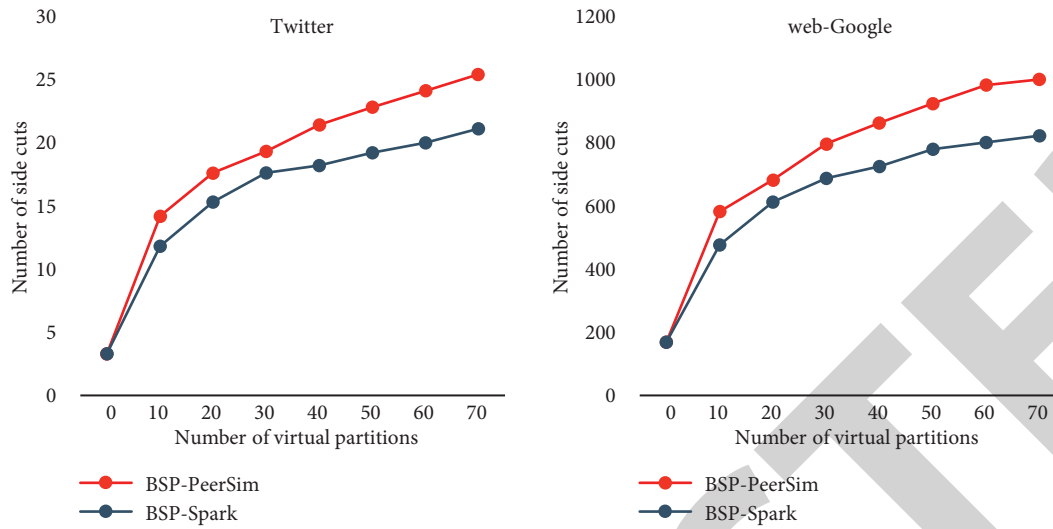


FIGURE 8: Relationship between the two partitions of Twitter and web-Google and the edge cut number of the JA-BE-JA algorithm.

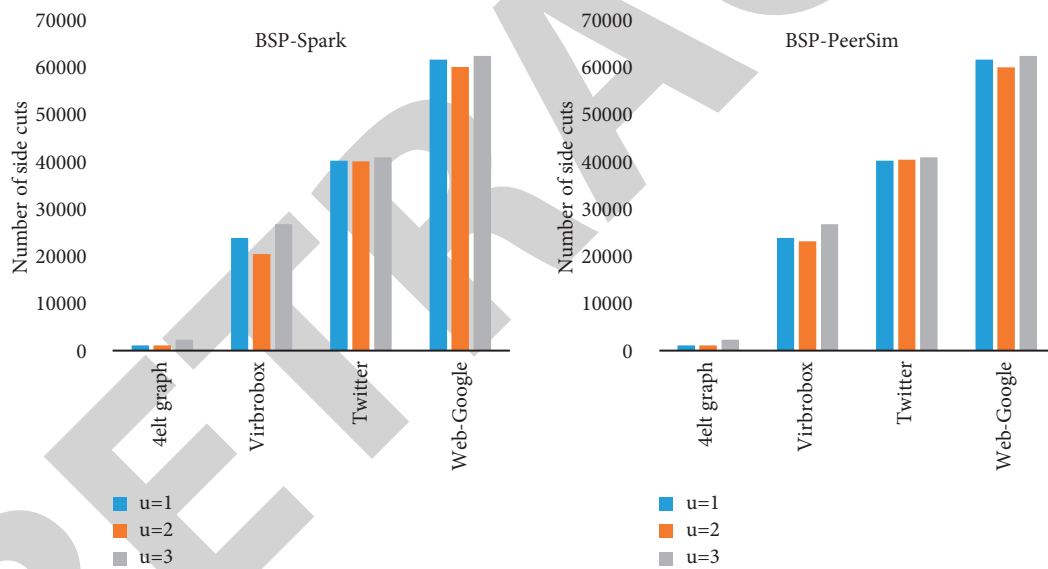


FIGURE 9: Relationship between  $u$  and JA-BE-JA algorithm in the original BSP-PeerSim and BSP-Spark environments.

It can be seen that the final reconstruction result of the distribution network is to close the tie switch between nodes 10 and 14, close the tie switch between 5 and 11, open the section switch between 8 and 10, turn on the segment switch between 9 and 11, and use the improved JA-BE-JA algorithm graph theory to divide the network loss to 454.3 KW. The initial network loss is 508.6 KW, and the network loss is reduced by 10.68%. The power distribution system is often of large-scale and widely distributed, and there are multiple feeders. It can be seen that this algorithm is still applicable in large-scale distribution networks, reducing the burden of power flow calculation, reducing the network loss, and also solving the problem of local optimization.

## 5. Discussion

This article focuses on the analysis of large-scale distribution network reconstruction with graph theory and graph partitioning algorithms. Image segmentation based on graph theory has been developed in many fields and has made indelible contributions in medicine and neurology [25]. This article first describes many advantages of graph theory and graph segmentation technology in the application of distribution networks and then describes the research methods of graph theory and graph segmentation and distribution networks. For example, the image segmentation method based on graph theory-full variational approximation of the Ginzburg-Landau functional of graph clustering uses

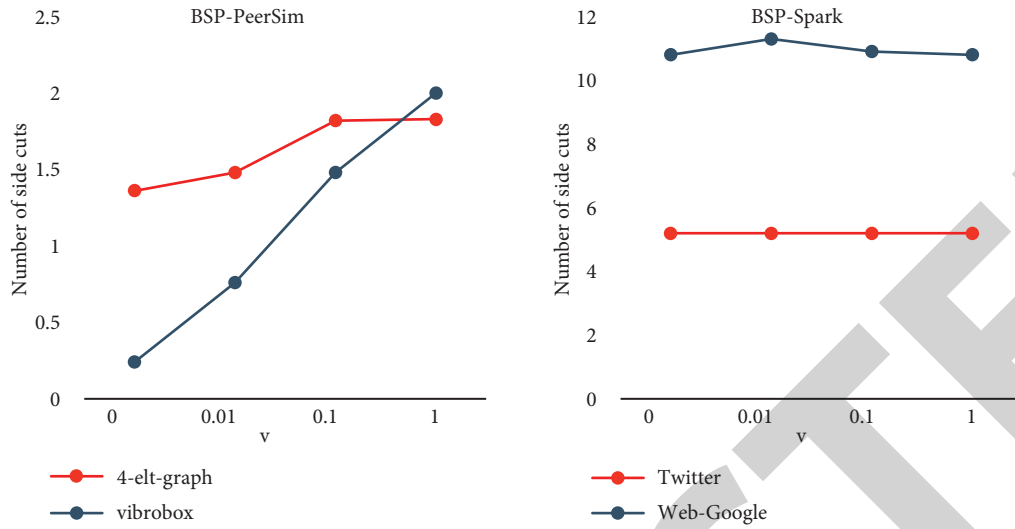


FIGURE 10: Relationship between  $v$  and the JA-BE-JA algorithm edge cut number.

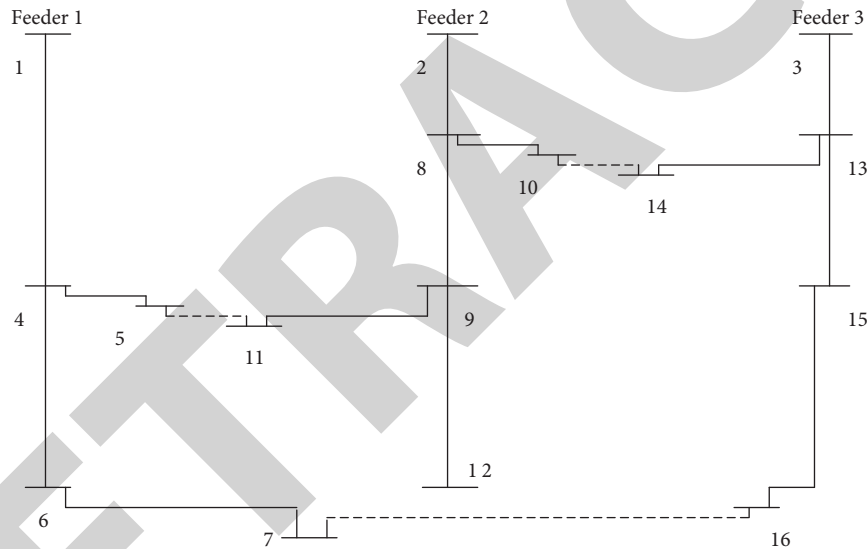


FIGURE 11: Bus distribution system.

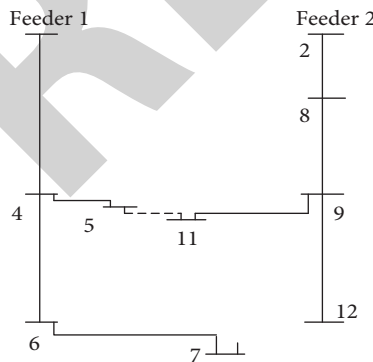


FIGURE 12: Subgraph after halving processing.

quantum mechanics, superconductor, electrodynamics, and other phenomena to derive relevant formulas, divides the feature vector through the clustering algorithm, and performs

image segmentation. For example, for the power flow algorithm of distribution network reconstruction, the radial circuit diagram is given in this article, and the power flow algorithm is used for calculation according to the topology of the distribution network. In this paper, the improved Niu pull method is used to improve convergence, timeliness, etc., phase angle correction and voltage correction are carried out, the distribution network reconstruction with DG is proposed, and the optimal value of the population particles is repeatedly iterated. The IEEE69 example is used to show the situation, and the forward and backward power flow calculations are used to iteratively converge to achieve the minimum network loss. It also describes the fault recovery model of the distribution network, using the application layer, decision-making layer, and control system of the distribution network to write the reconstruction algorithm into the distribution network dispatching system. The faults of the distribution network are quickly located and isolated through the FTU, and quickly

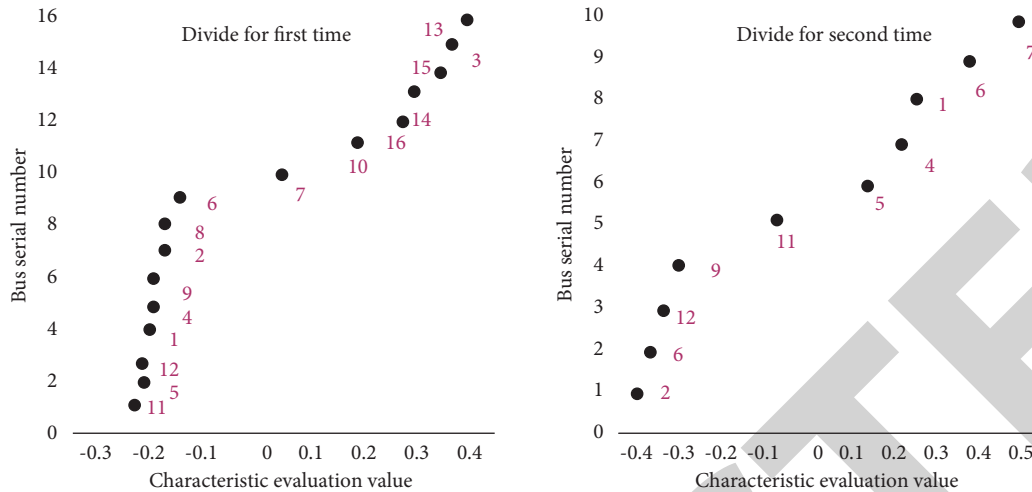


FIGURE 13: Comparison of feature evaluation values after the first and second sorting.

repaired. Finally, the graph theory algorithm and the performance of the improved JA-BE-JA algorithm are tested through experiments, and they are applied to the distribution network, and the relationship between each partition and the JA-BE-JA algorithm is compared. Substituting the parameters into the test, it can be seen that a 4-way partition will cause migration of a large number of vertex data, and migration of only a small part of the data to different partitions. Finally, a 16-bus distribution system example is used for testing, and the improved JA-BE-JA algorithm can reduce the loss by 10.68% compared with the previous one, and it can be applied to a multifeeder system, that is, the large-scale distribution network, which reduces the power flow calculation. It can also reduce network loss and prevent local optimization. It can be obtained that this improved algorithm can be better applied in large-scale distribution networks.

## 6. Conclusions

In the analysis of large-scale distribution network reconstruction with graph theory and graph partitioning algorithms, it is known that based on the BSP-Spark environment, the improved algorithm has robust stability and can avoid the problem of local optimization. A 4-way partition caused 76.3% of the vertex data migration, and only 23.7% moved to a different partition. The improved JA-BE-JA algorithm is also more obvious for the optimization of the distribution network. Using the improved JA-BE-JA algorithm with graph theory and graph division, the network loss is 454.3 KW, the initial network loss is 508.6 KW, and the network loss is reduced by 10.68%. In addition, the improved algorithm can effectively increase the number of edge cuts and the number of vertices exchanged on the distributed clusters. It is much more effective than the traditional graph segmentation algorithm in the processing of large-scale distribution networks.

## Data Availability

The data used to support the findings of this study are available from the author upon request.

## Conflicts of Interest

The author declares no conflicts of interest.

## References

- [1] X. J. Li and G. H. Yang, "Graph theory-based pinning synchronization of stochastic complex dynamical networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 2, pp. 1–11, 2017.
- [2] C. L. Keown, M. C. Datko, and C. P. Chen, "Network organization is globally atypical in autism: a graph theory study of intrinsic functional connectivity," *Biological Psychiatry Cognitive Neuroscience & Neuroimaging*, vol. 2, no. 1, pp. 66–75, 2017.
- [3] F. Glaria, C. Hernández, and S. Ladra, "Compact structure for sparse undirected graphs based on a clique graph partition," *Information Sciences*, vol. 544, no. 7043, pp. 485–499, 2021.
- [4] Y. Li, Y. Zhao, and Q. Wu, "A graph-based thread partition approach in speculativemultithreading," *Concurrency: Practice and Experience*, vol. 29, no. 21, pp. e4294.1–e4294.22, 2017.
- [5] S. Ghasemi, "Balanced and unbalanced distribution networks reconfiguration considering reliability indices," *Ain Shams Engineering Journal*, vol. 9, no. 4, pp. 1567–1579, 2018.
- [6] H. Hong, Z. Hu, R. Guo, and M. Jun, "Directed graph-based distribution network reconfiguration for operation mode adjustment and service restoration considering distributed generation," *Journal of Modern Power Systems & Clean Energy*, vol. 5, no. 1, pp. 142–149, 2017.
- [7] T. Ding, C. Li, C. Zhao, and M. Wang, "Total supply capability considering distribution network reconfiguration under Nk transformer contingency and the decomposition method," *IET Generation, Transmission & Distribution*, vol. 11, no. 5, pp. 1212–1222, 2017.
- [8] K. Teeparthi and D. V. Kumar, "Multi-objective hybrid PSO-APO algorithm based security constrained optimal power flow with wind and thermal generators," *Engineering Science & Technology An International Journal*, vol. 20, no. 2, pp. 411–426, 2017.
- [9] A. MD, B. HC, and C. NL, "Dynamic population artificial bee colony algorithm for multi-objective optimal power flow," *Saudi Journal of Biological Sciences*, vol. 24, no. 3, pp. 703–710, 2017.



## Retraction

# Retracted: An Algorithm for Time Prediction Signal Interference Detection Based on the LSTM-SVM Model

### Computational Intelligence and Neuroscience

Received 1 August 2023; Accepted 1 August 2023; Published 2 August 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] N. Xiao and Z. Song, "An Algorithm for Time Prediction Signal Interference Detection Based on the LSTM-SVM Model," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 1626458, 8 pages, 2022.

## Research Article

# An Algorithm for Time Prediction Signal Interference Detection Based on the LSTM-SVM Model

Ningbo Xiao  and Zuxun Song 

*School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710072, Shaanxi, China*

Correspondence should be addressed to Ningbo Xiao; [xiaoningbo@bjwlxy.edu.cn](mailto:xiaoningbo@bjwlxy.edu.cn)

Received 14 January 2022; Accepted 17 February 2022; Published 11 March 2022

Academic Editor: Daqing Gong

Copyright © 2022 Ningbo Xiao and Zuxun Song. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Interference detection is an important part of the electronic defense system. It is difficult to detect interference with the traditional method of extracting characteristic parameters for interference generated at the same frequency as the original signal. Aiming at this special time-frequency overlapping interference signal, this paper proposes an interference detection algorithm based on the long short-term memory-support vector machines (LSTM-SVM) model. LSTM is used for the time series prediction of the received signal. The difference between the predicted signal and the received signal is used as the feature sample, and the SVM algorithm is used to classify the feature samples to obtain the recognition rate of whether the sample has interference. The LSTM-SVM model is compared with the gate recurrent unit-support vector machines (GRU-SVM) model, and the comparison results are visualized using a confusion matrix. The simulation results show that this LSTM-SVM model algorithm cannot only detect the existence of the interference signal but also can determine the specific position of the interference signal in the received waveform, and the detection performance is better than the GRU-SVM model.

## 1. Introduction

In the present study, algorithms based on deep learning are booming, and research on algorithms based on deep learning has been proposed in many fields. LSTM neural networks can better predict time series data, and various algorithms based on LSTM to predict future development are also being widely used. Natural disaster prediction is a very important application. LSTM is used to learn the spatiotemporal relationship between disasters in different locations, and this relationship is used to make predictions. Through the training of historical trajectory data, the trajectory prediction model is obtained, so that early warning can be achieved [1, 2]. In terms of health care, the use of an LSTM network to independently detect the contraction curve of an ECG or EMG signal, based on the predicted signal, provides a powerful assistant for medical staff to quickly diagnose and help patients recover [3–5]. Adopting LSTM can achieve higher accuracy than traditional mathematical models in predicting future disease trends [6]. In

terms of traffic prediction, new methods for urban traffic prediction combining spectrogram analysis and LSTM have been proposed [7, 8]. Using an LSTM network, electric cars can estimate future speed in real-time to estimate the precise value of brake pressure [9]. In addition, subway passenger flow prediction using LSTM can guide passengers to effectively select departure times and transfer from station to station [10]. It is widely used in fault detection, such as image forgery detection, analyzing discriminative features between different regions by combining encoder and LSTM networks, to predict whether the image has been tampered with [11]. LSTM can provide inpainting for video frames, using convolutional LSTM-based encoders and decoders to predict missing frames [12]. LSTM can achieve a predicted quantified output of the remaining mechanical life [13].

At present, the related research involving interference detection in wireless communication is increasing day by day, and some effective algorithms have been proposed. Some of these algorithms are based on computer network protocols and detect the presence of interference signals by

detecting the listening time of the carrier signal; Some algorithms analyze the signal according to the frequency domain transformation, and perform the frequency domain transformation on the signal to extract features, then analyze whether there is interference in the signal; some algorithms achieve the purpose of signal analysis by counting the mathematical characteristics of the signal and calculating the high-order statistics of the signal so as to extract the signal characteristics. Obviously, the above method belongs to the traditional interference detection method, which mainly analyzes the parameter characteristics such as the spectrum and power of the received signal, and then confirms whether there is interference. Secondly, when the signal overlaps with the interfering frequency, the eigenvalues of the interfering signal are difficult to detect [14].

Since communication signals belong to time series data, the time series prediction algorithm based on deep learning provides an idea for signal interference detection. Among them, the recurrent neural network has the characteristics of short-term memory. If the time series signal is long, it will affect the transmission of information. LSTM and gate recurrent unit (GRU) can solve short-term memory problems. Based on the existing research field, we combine LSTM and GRU with the SVM model and propose the LSTM-SVM and GRU-SVM models for signal interference detection [15].

## 2. Materials and Methods

A signal is a periodic time series data, and unknown sampling points can be predicted through the correlation between sampling points [16]. According to this feature, assuming that the training signals are all normal signals, the predicted future signals should also be normal signals. In this case, calculating the difference between the real signal and the predicted signal as the detection feature can get a good detection effect. The signal prediction problem can be solved by the LSTM network. In essence, signal prediction is a time series prediction problem with periodic characteristics. The previously received signal is used as the input, and the output is the signal at the moment you want to predict.

**2.1. Structure of LSTM.** The LSTM unit structure is shown in Figure 1. Compared with the RNN cell structure, LSTM adds a cell state to the structure. LSTM neurons consist of a forget gate, an input gate, and an output gate. The forget gate is used to determine the importance of information, the input gate is used to update the cell state, and the output gate is used to calculate the value of the hidden state.

**2.2. Structure of GRU.** The GRU unit structure is shown in Figure 2. Compared with the LSTM cell structure, the GRU reduces the cell state. The neurons of the GRU include an update gate and a reset gate [17]. The update gate is used to determine the importance of the information, while the reset gate is used to determine the proportion of retained information. Compared to LSTM, GRU has simpler neurons and therefore faster training.

**2.3. Activation Function: Tanh, Sigmoid.** The Tanh function is used to adjust the threshold of the function output. The value range is  $[-1, 1]$ , as shown in Figure 3.

$$f(x) = \frac{\sinh(x)}{\cosh(x)} = \frac{1 - e^{-2x}}{1 + e^{-2x}} \quad (1)$$

The activation function Sigmoid is similar to the tanh function, with a value range of  $(0, 1)$ , as shown in Figure 4.

$$\sigma = \frac{1}{1 + e^{-x}} \quad (2)$$

**2.4. Time Series Signal Forecasting.** 60% of the received signal is used for training datasets of LSTM and GRU. The remaining 40% of the signals, as the test dataset, are used for prediction. The eigenvalues are obtained by subtracting the test data from the predicted data. The process is shown in Figure 5 [18].

When the normal communication signal in the received signal adopts the cosine signal and the interference signal adopts the narrow-band Gaussian noise [19], the LSTM time series signal prediction is shown in Figure 6, the obtained eigenvalues are shown in Figure 7.

The first picture of Figure 6 is the original received signal. You can see that there are interference signals between 2500–3000. The second picture is the training and prediction time series of LSTM, where the first 60% is the training set, and the last 40% is the prediction value. The first picture in Figure 7 is the 40% of the received signal that needs to be predicted. The second picture is the comparison of the predicted value with the original received signal, and the third simulation plot is the eigenvalues [20].

**2.5. Windowing of Signal Sample Values.** In the detection process, when the deviation of each sampling point is used as a feature to train the SVM classifier, the difference calculated for the test dataset, especially the part containing the interference signal, contains many differences similar to the difference calculated from the normal signal points, and these points are difficult to detect. So, we add windows to the features based on the time step and detect them based on the waveforms and correlations of multiple feature points to overcome contingency. The windowing process is shown in Figure 8.

If only each sampling point were used as a feature of interference detection, the detection effect would be very poor due to the accidental error of calculating the difference. Therefore, here we select the mean value of the sampling points of each  $n$  continuous signal as a feature, and the method of constant translation is used to construct a complete feature of interference detection. As shown in Figure 8, the signal is windowed, the window length is “ $N$ ,” the step length is “ $S$ ,” and the input data is formed in time series. The value of  $S$  is 1, when  $N$  is an odd number, the  $i$ th eigenvalue is calculated according to formula (3), and when  $N$  is an even number, the  $i$ th eigenvalue is calculated according to formula (4).

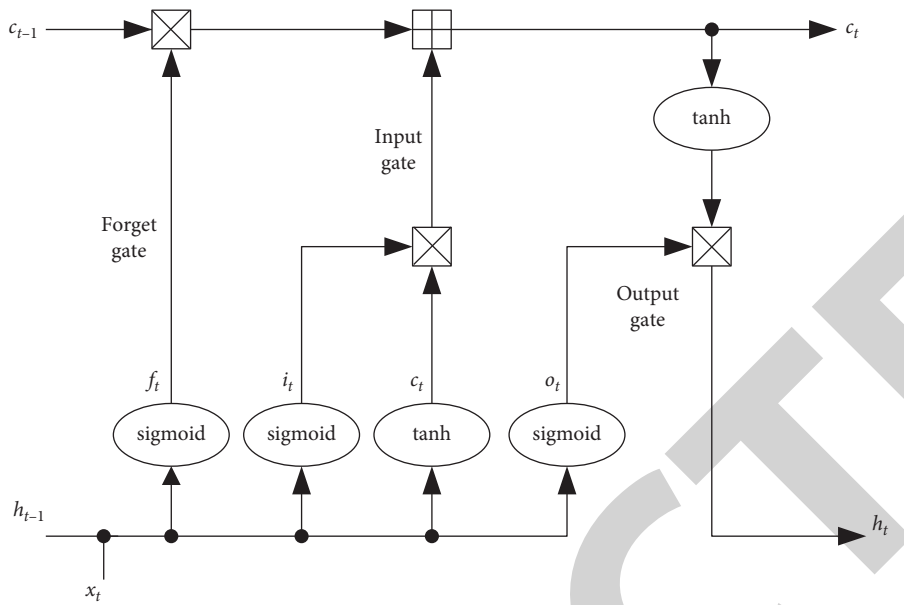


FIGURE 1: Structure of LSTM.

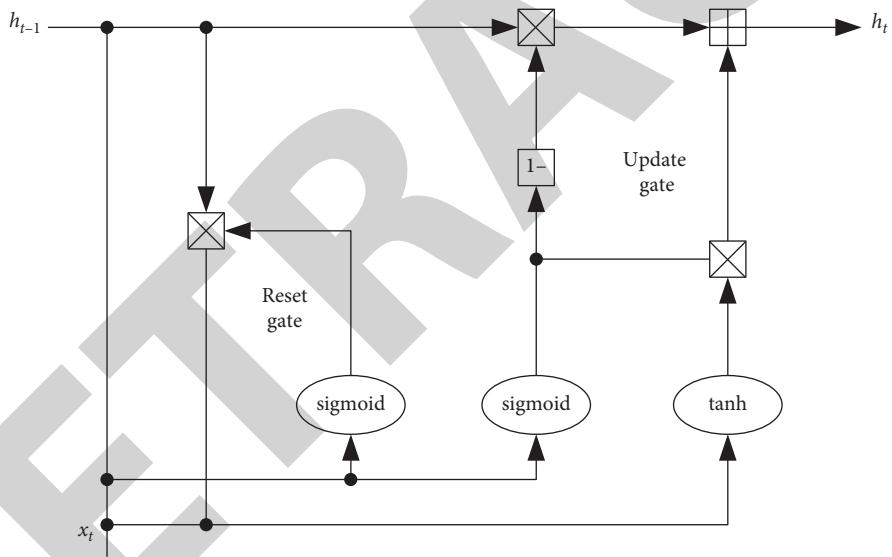


FIGURE 2: Structure of GRU.

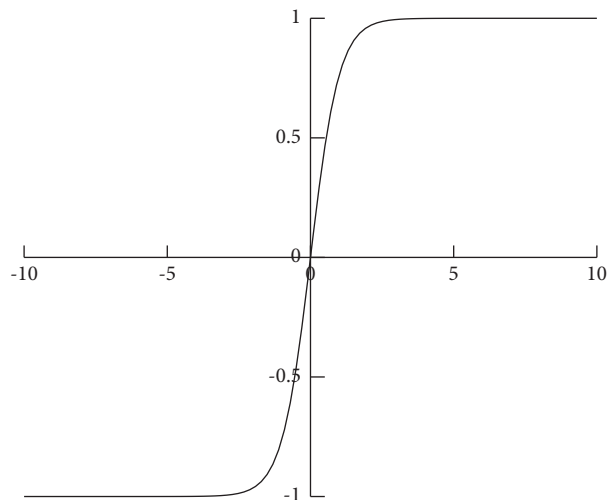


FIGURE 3: Activation function tanh.

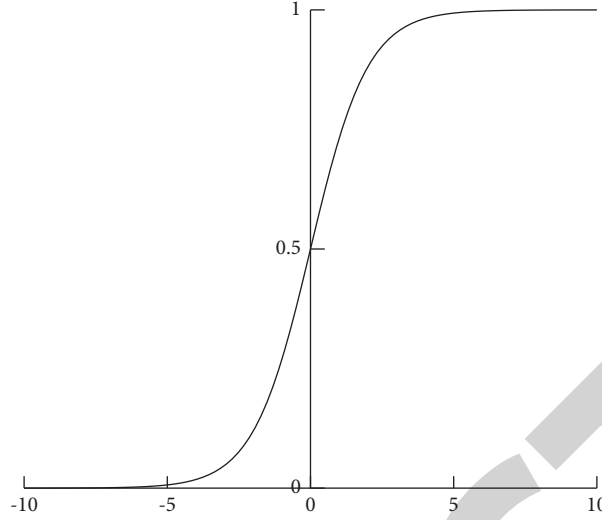


FIGURE 4: Activation function sigmoid.

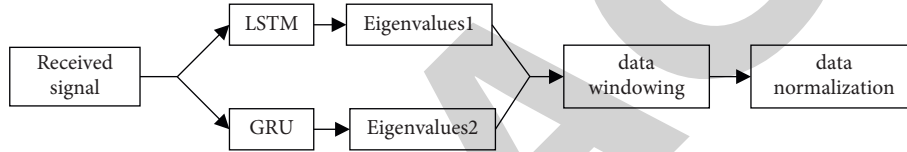


FIGURE 5: Block diagram of the signal prediction process.

$$V_{i-e} = \frac{V_{i-(N-1)/2} + \dots + V_{i-1} + V_i + V_{i+1} + \dots + V_{i+(N-1)/2}}{N} \quad (3)$$

$$V_{i-e} = \frac{V_{i-N/2+1} + \dots + V_i + V_{i+1} + \dots + V_{i+N/2}}{N} \quad (4)$$

**2.6. Data Normalization Processing.** The eigenvalues are normalized and preprocessed, the normalization method adopted is according to the following formula:

$$f: x \longrightarrow y = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (5)$$

### 3. Results

In the AWGN channel environment, the sampling frequency of the received signal is 3 KHz, the center frequency of the communication signal and the interference signal is 1 KHz, and the signal power is 10 dB. The sampled signal of the received signal  $r(t)$  is  $r(n)$ , as shown in formula (6).

$$r(n) = s(n) + j(n) + w(n), n = 1, 2, \dots, N \quad (6)$$

In the above equation  $N$  is the sampling length,  $s(n)$  is the communication signal,  $j(n)$  is the interference signal, and  $w(n)$  is the Gaussian white noise.

**3.1. Comparison of the Influence of SNR on Interference Detection.** Both LSTM and GUR are set to train the first 0.6 parts of the time series of the data and test the remaining 0.4

parts at 100 training iterations. After obtaining the eigenvalues, the eigenvalues are not subjected to windowing processing, and the eigenvalues are classified in the SVM to obtain the interference detection rate. When the signal power is 10 dB, the SNR value range is 0dB–20 dB, and the step size is 2 dB, the interference detection accuracy is shown in Figure 9.

It can be seen from Figure 9 that, the difference data obtained by the LSTM and GUR time prediction models are not windowed. As the SNR increases, the detection rate of the difference signal by SVM also gradually increases.

**3.2. Comparison of Windowing Experiments.** Due to the continuity characteristics of the interference signal, taking into account the correlation between the sampling points, window processing is performed on the sampling points, which can overcome the influence of accidental errors. When the power of the signal is 10 dB and the SNR is 20 dB, the eigenvalues are processed by windowing. The window length is 1–10, the step size is 1, and the interference detection accuracy is obtained as shown in Figure 10.

It can be seen from Figure 10 that LSTM has the highest detection rate when the window length is 6, but the GRU windowing effect is not good. The detection rate is the highest when no window is added, so the window value is selected as 1.

**3.3. Interference Detection Accuracy under the Optimal Window.** The difference signal generated by the LSTM is

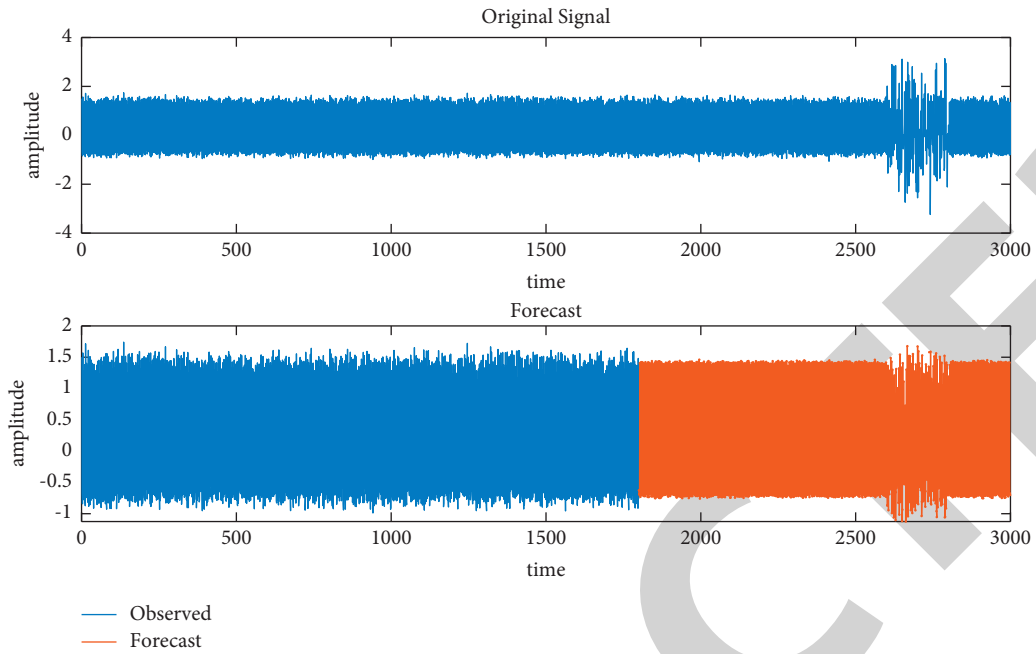


FIGURE 6: LSTM timing signal prediction.

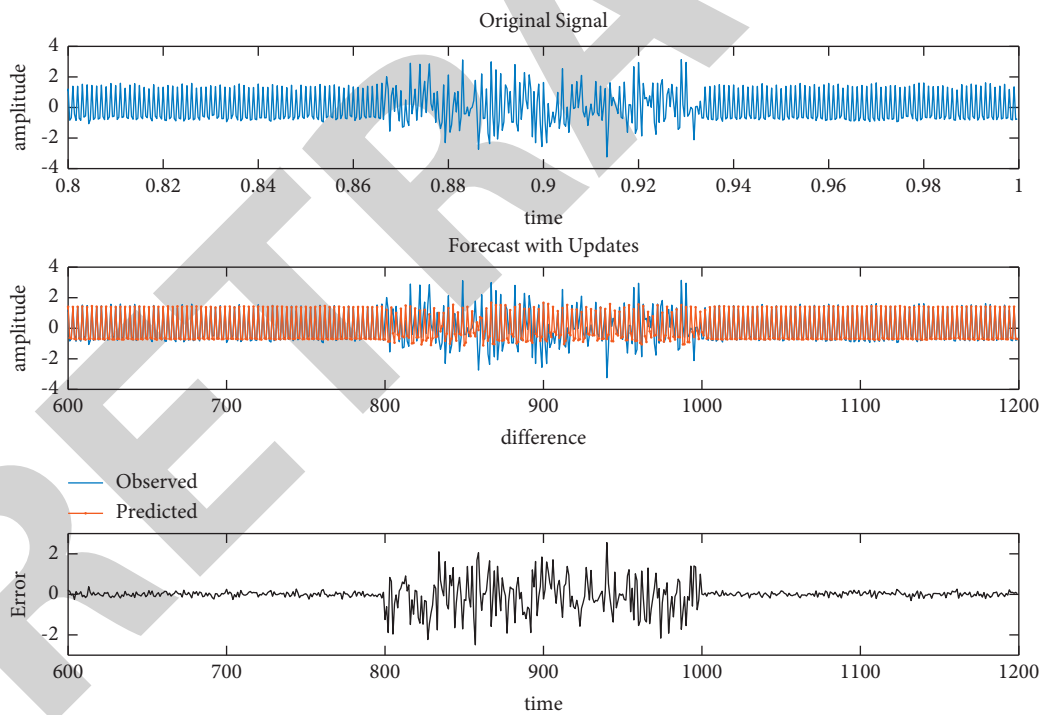


FIGURE 7: Eigenvalues.

windowed with a length of 6 and the difference signal generated by the GRU is not windowed. After normalization, it is sent to the SVM for classification. The classification result diagram is shown in Figure 11 [21].

It can be seen from Figure 11 that, when there is an interference signal, the interference detection accuracy of LSTM-SVM is 93.6%, and the interference detection accuracy of GRU-SVM is 89.6%.

**3.4. Confusion Matrix.** Since the SVM classification results cannot see the detection accuracy of each type, we use a confusion matrix to further refine the detection results. The confusion matrices for LSTM-SVM and GRU-SVM interference detection classification are shown in Figure 12 [22].

From Figure 12, it can be seen that, when there is no interference, the detection accuracy of the GRU-SVM is 100%, and that the detection accuracy of the LSTM-SVM



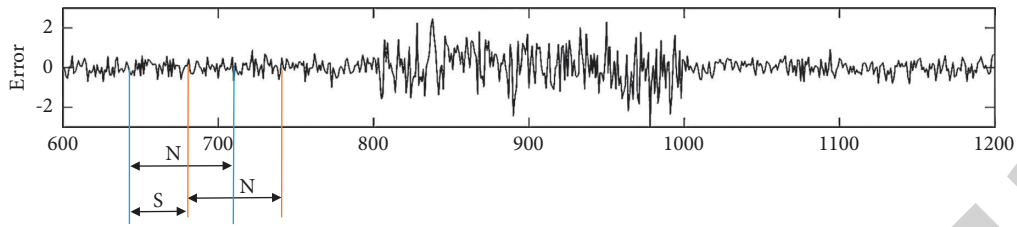


FIGURE 8: Signal data windowing process.

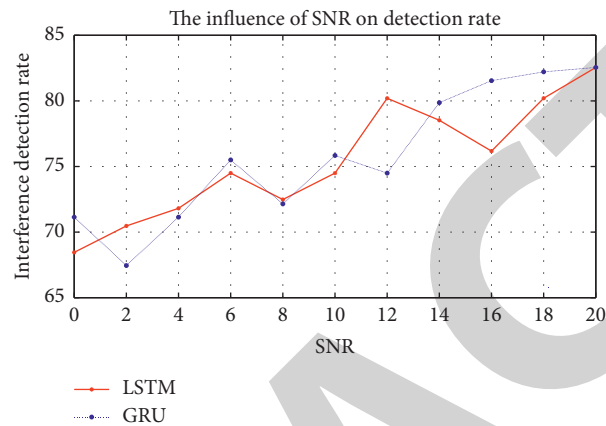


FIGURE 9: Influence of SNR on detection accuracy.

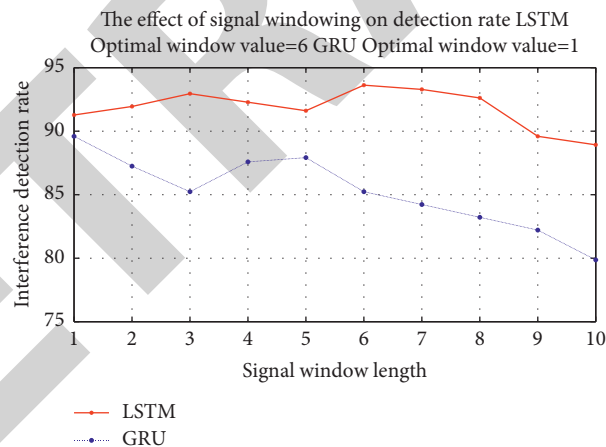


FIGURE 10: Effect of signal windowing on detection rate.

model is 99.5%. When there is interference, the detection accuracy of GRU-SVM is 68.7%, and the detection accuracy of LSTM-SVM is 81.8%. Overall, the false detection rate of GRU-SVM is 10.4%, and the false detection rate of LSTM-SVM is 6.4%. Therefore, the eigenvalues obtained by LSTM-SVM are more conducive to detecting the presence of interference signals.

#### 4. Discussion

The traditional interference detection starts from the received signal, analyzes the frequency domain characteristics of the received signal, extracts the characteristic value, and

then determines whether there is an interference signal. However, when the interference signal and the communication signal have the same frequency, it is difficult to extract the eigenvalues by using the traditional method. From the perspective of the time domain, this paper adopts the LSTM model, which has good prediction ability for time series signals. The eigenvalues are obtained from the difference between the predicted signal and the received signal, which solves the problem of signal interference detection in the case of interference.

The setting of the interference signal in the experiment described in this paper is relatively ideal. The interference signal used in this paper is a partial frequency band

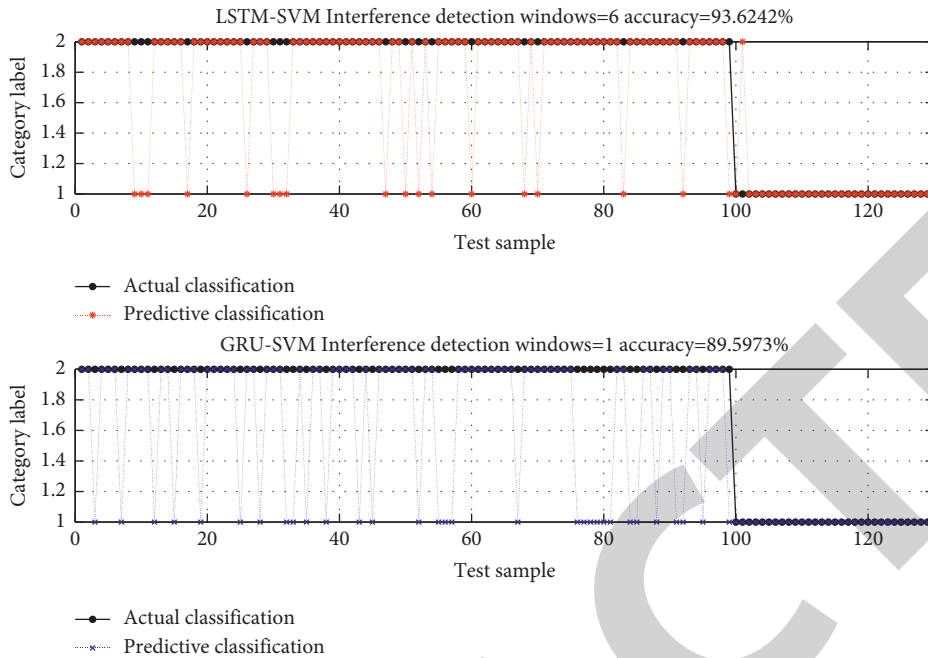


FIGURE 11: Simulation of interference detection accuracy.

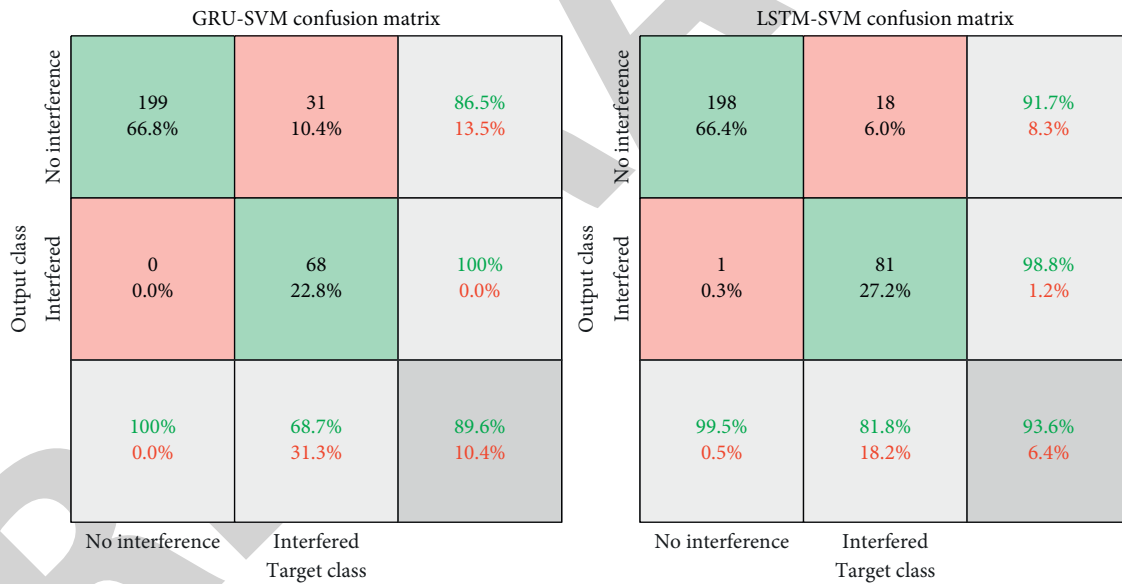


FIGURE 12: Confusion matrix of LSVM-SVM and GRU-SVM classification.

interference signal, that is, the interference signal is generated by Gaussian noise through a narrow-band filter. You can try to use other interference signals to verify whether the experimental results in this paper are applicable.






### 5. Conclusion

Aiming at the detection of special time-frequency overlapping interference signals, an interference detection algorithm based on the LSTM-SVM model is proposed in this paper. LSTM is used for the time series prediction of received signals, and the difference between the predicted

signal and the received signal is used as a feature sample. This method can still be trained using a small amount of data and can obtain better prediction performance. The eigenvalue samples are windowed, and the SVM algorithm is used to classify and detect the eigenvalue samples. The LSTM-SVM model is compared with the GRU-SVM model. The comparison results show that the interference detection accuracy of the two models increases with the improvement of the SNR. Windowing the eigenvalues can significantly improve the interference detection performance of the LSTM-SVM model, but windowing the GRU-SVM model will reduce the accuracy. Under the optimal window, the detection

## Research Article

# Carbon Transfer Decision Model Based on LMDI Method

Yawei Qi , Guangping Rao , Lei Zha , Lu Chen , and Yuping Niu 

*School of Information Management, Jiangxi University of Finance and Economics, Nanchang 330032, China*

Correspondence should be addressed to Yawei Qi; [qiyawei@jxufe.edu.cn](mailto:qiyawei@jxufe.edu.cn)

Received 10 January 2022; Accepted 17 February 2022; Published 9 March 2022

Academic Editor: Daqing Gong

Copyright © 2022 Yawei Qi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Establishing a coordinated governance mechanism for regional carbon emissions is an essential way to achieve carbon peak and carbon neutrality, while the study of interprovincial carbon emissions transfer is one of the important foundations of regional carbon emissions coordinated governance research. Based on the multiregional input-output (MRIO) model, this study calculated the carbon emissions from both the producers' perspective and the consumers' perspective and analyzed the interprovincial net carbon emissions transfer decision. Furthermore, the logarithmic mean Divisia index (LMDI) method was adopted to decompose the factors that affect the province's net carbon emissions into technological effect, structural effect, input-output effect, and scale effect. It was revealed that the input-output effect was the primary influencing factor of the net carbon transfer at the provincial level.

## 1. Introduction

China's economy has grown massively since the 21st century. As of 2020, China has achieved its first centenary goal—building a moderately prosperous society in all respects and is striding forward to its second centenary goal. However, the emission of carbon dioxide (CO<sub>2</sub>) is also rising along with the total economic volume, and China has become the world's largest carbon emitter. The issue of greenhouse gas emissions is getting more and more attention. On March 15, 2021, the Chinese government stated to have carbon emissions peak before 2030 and achieve carbon neutrality by 2060. This is a major decision under the goal of building a community of shared future for all humankind.

China has a vast territory. The differences in resource endowments and climate environment in different regions have not only brought differences in industrial structure between regions but also differences in CO<sub>2</sub> emissions of regional production activities. During regional trading, the production of goods and the provision of services lead to carbon emissions in production areas, but goods and services are actually consumed by consumption areas; therefore, CO<sub>2</sub> emissions are transferred along with the flow of goods and services. If the inter-regional transfer of carbon emissions is ignored during

the promotion of carbon emissions reduction, on the one hand, it may cause some regions to import high-carbon products from other regions for carbon emissions reduction. It will impair the effectiveness of emissions reduction policy, and even cause partial decrease with overall increase [1, 2]. On the other hand, it may cause high-polluting and high-energy-consuming industries to migrate from high-carbon emissions areas to low-carbon emissions areas. It is not conducive to take advantage of resource endowments in various regions, and neither conducive to promoting the development and utilization of clean energy and encouraging low-carbon consumption and low-carbon lifestyle. Therefore, to advance carbon emissions reduction, we need to think from the perspective of regional synergy. However, before planning how to scientifically promote the fair distribution of CO<sub>2</sub> reduction between regions, we need to pay attention to the following two issues. First, how is the regional distribution of carbon emissions transfer in China? Second, what are the main factors affecting the transfer of carbon emissions between regions in China? Therefore, this study is committed to carrying out the analysis of interprovincial carbon emission transfer, characterizing the relationship of carbon emission transfer, discussing its influencing factors, and consolidating the foundation for the research on carbon emission collaborative governance.

## 2. Literature Review

With respect to the calculation of carbon emissions transfer, more and more scholars adopted the multiregional input-output (MRIO) model in recent years. The single-region input-output (SRIO) model was first constructed by Leontief [3] and then expanded to multiple regions and widely applied in environmental impact assessment (EIA) [1, 2, 4].

Various scholars have integrated the MRIO model with the actual situation in China. Guo et al. [5], Yao et al. [6], and Zhang et al. [7] used the MRIO model to calculate and analyze China's interprovincial carbon emissions input-output, the carbon footprint of residents' consumption in China's eight major regions, and the carbon emissions input-output from various sectors in the eight major regions, respectively, revealing significant regional differences [5–7]. In addition, Liu et al. [8] used the 1997 inter-regional input-output table to analyze the transfer of carbon emissions from the perspective of industry and proposed a CO<sub>2</sub> emissions reduction model for industrial structure adjustment. The calculation demonstrated that structural adjustment of energy consumption of electric power industry and improving the efficiency of energy utilization in heavy industry were effective means to reduce carbon emissions [8]. Su and Ang [9, 10] explored the spatial agglomeration of carbon emissions transfer by calculating the carbon emissions transfer between major regions in China in 2002 and proposed to optimize interregional carbon emissions transfer from the mechanism level [9, 10]. Sun et al. [11] calculated carbon transfers between Indian economic sectors from 1995 to 2009 based on the SRIO model. The study found that CON (construction) transfers the largest amount, and EGW (electricity, gas, water supply) emissions account for 60% of CO<sub>2</sub> in the secondary industry [11]. Based on the MRIO model, Chen et al. [12] independently compiled China's inter-regional input-output table for 2012, calculated the amount of inter-provincial carbon emissions transfer, and analyzed the inter-provincial carbon equity from the perspective of carbon Gini coefficient [12]. Wang and Hu [13] used an inter-regional bilateral trade carbon transfer measurement model to measure the carbon emissions caused by interprovincial demand and interprovincial exports in China in 2007, 2010, and 2012. The study shows that cooperation and diffusion of emission reduction technologies can be effective suppress interprovincial carbon emissions transfer [13]. Han et al. [14] explored the flow of carbon emissions in China's provinces and other economies from a multi-regional perspective based on Nested IOA [14]. Based on the MRIO model, Wang et al. [15] analyzed the carbon-neutrality-oriented carbon emissions reduction model by studying the consumption-based carbon emissions and carbon transfer at the provincial and industry levels in China [15]. Liu et al. [16] deduced the calculation method of the value chain embodied carbon transfer based on the input-output table balance formula, and found that China's inter-regional value chain embodied carbon transfer and its net value both showed an increasing trend [16].

Regarding the decomposition of the influencing factors of carbon emissions transfer, the popular academic methods include the structural decomposition and the logarithmic

mean Divisia index (LMDI) decomposition, both of which have been widely used in various research scenarios. Based on the input-output model, Qian and Yang [17], as well as Jiang [18], employed the structural decomposition method to decompose carbon emissions embodied in international trade between East Asia and the BRICS, respectively [17, 18]. Guo [19], Du and Sun [20] and Huang et al. [21] applied the LMDI method to decompose the total carbon emissions, export embodied carbon emissions, and carbon emissions embodied in inter-provincial trade [19–21].

Existing studies have adequately applied the MRIO model to calculate the international and domestic carbon emissions transfer, used different methods to decompose the influencing factors, and analyzed interprovincial carbon equity from various perspectives and put forward corresponding proposals of carbon emissions reduction. On these bases, this study used a MRIO model to calculate the carbon emissions transfer between provinces based on the three-year data of 2012, 2015, and 2017, analyzed the transfer paths of inter-regional carbon emissions through social network, and then decomposed the factors affecting the net carbon emissions transfer at the provincial level using the LMDI method. We aimed to provide a basis for the construction of a coordinated governance mechanism for regional carbon emissions and emissions reduction plans to promote the achievement of carbon peak and carbon neutrality.

## 3. Interprovincial Carbon Emissions Transfer

Different provinces and cities in China have different resource endowments and natural environments. When designing emissions reduction plans, it will be not all-inclusive to consider only the differences in industrial structure between provinces and cities caused by different resource endowments. During the implementation of the emission reduction plans, because interprovincial trade will drive the transfer of interprovincial carbon emissions, with no restriction, some regions will purchase high-carbon emission products from other regions to replace local production for the purpose of carbon reduction, sabotaging overall emissions reduction. Due to differences in technological levels between regions, there may even be reduced partial emissions with increased overall emissions, the so called inter-regional carbon leakage. For that reason, we attempted to provide a basis for the coordinated governance mechanism of regional carbon emissions by calculating and analyzing the current interprovincial carbon emissions transfer, carbon transfer network, and carbon transfer path.

*3.1. Interprovincial Carbon Emissions Transfer Calculation Model.* In the present study, a MRIO model was adopted to calculate the carbon emissions from both the consumers' perspective and the producers' perspective and the inter-region carbon emissions transfers between 31 provinces and cities across China.

Assuming that in the MRIO model there are totally  $n$  provinces and cities and  $m$  industries in each province/city, the MRIO model has the following equation:

$$\begin{bmatrix} X_1^1 \\ X_2^1 \\ \vdots \\ X_m^n \end{bmatrix} = \begin{bmatrix} X_{1,1}^{1,1} + X_{1,2}^{1,1} + \cdots + X_{1,m}^{1,n} \\ X_{2,1}^{1,1} + X_{2,2}^{1,1} \cdots + X_{2,m}^{1,n} \\ \vdots \\ X_{m,1}^{n,1} + X_{m,2}^{n,1} + \cdots + X_{m,m}^{n,n} \end{bmatrix} + \begin{bmatrix} Y_1^{1,1} + Y_1^{1,2} + \cdots + Y_1^{1,n} \\ Y_2^{1,1} + Y_2^{1,2} + \cdots + Y_2^{1,n} \\ \vdots \\ Y_m^{n,1} + Y_m^{n,2} + \cdots + Y_m^{n,n} \end{bmatrix} + \begin{bmatrix} EX_1^1 \\ EX_2^1 \\ \vdots \\ EX_m^n \end{bmatrix}, \quad (1)$$

where  $X_i^I$  represents the total output of the industry  $i$  of province  $I$ ,  $X_{i,j}^{I,J}$  represents the intermediate input of the industry  $i$  of province  $I$  to the industry  $j$  of province  $J$ ,  $Y_i^{I,J}$

represents the final use of the industry  $i$  of province  $I$  by province  $J$ , and  $EX_i^I$  represents the output of the industry  $i$  of province  $I$ . Transform the intermediate input to get:

$$\begin{bmatrix} X_1^1 \\ X_2^1 \\ \vdots \\ X_m^n \end{bmatrix} = \begin{bmatrix} A_{1,1}^{1,1} & A_{1,2}^{1,1} & \cdots & A_{1,m}^{1,n} \\ A_{2,1}^{1,1} & A_{2,2}^{1,1} & \cdots & A_{2,m}^{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m,1}^{n,1} & A_{m,2}^{n,1} & \cdots & A_{m,m}^{n,n} \end{bmatrix} \begin{bmatrix} X_1^1 \\ X_2^1 \\ \vdots \\ X_m^n \end{bmatrix} + \begin{bmatrix} Y_1^{1,1} + Y_1^{1,2} \cdots + Y_1^{1,n} \\ Y_2^{1,1} + Y_2^{1,2} \cdots + Y_2^{1,n} \\ \vdots \\ Y_m^{n,1} + Y_m^{n,2} \cdots + Y_m^{n,n} \end{bmatrix} + \begin{bmatrix} EX_1^1 \\ EX_2^1 \\ \vdots \\ EX_m^n \end{bmatrix}, \quad (2)$$

where  $A_{i,j}^{I,J} = X_{i,j}^{I,J} / \sum_{k=1}^n \sum_{l=1}^m X_{l,j}^{k,J}$  represents the products and services of the industry  $i$  of province  $I$  directly consumed by per unit of output of the industry  $j$  of province  $J$ .

The matrix form of equation (2) is

$$X = AX + Y + EX, \quad (3)$$

where  $X$  is the total output vector ( $mn \times 1$ ),  $A$  is the direct consumption coefficient matrix ( $mn \times mn$ ),  $Y$  is the final use vector ( $mn \times 1$ ), and  $EX$  is the export vector ( $mn \times 1$ ).

Transform equation (3) to get

$$X = (I - A)^{-1} (Y + EX). \quad (4)$$

Let  $B = (I - A)^{-1}$ , where  $B$  is the total input consumption coefficient matrix and the element  $B_{i,j}^{I,J}$  represents the intermediate products and final products provided by the industry  $i$  of province  $I$  to meet the per unit of final use in the industry  $j$  of province  $J$ .

Following the Intergovernmental Panel on Climate Change (IPCC) guidelines, this study calculates the carbon emissions of each industry in each province/city using the following equation:

$$\begin{aligned} C_i^I &= \sum_k CE_{ik}^I \\ &= \sum_k AD_{ik}^I \times NCV_k^I \times CC_i^I \times O_{ik}^I, \end{aligned} \quad (5)$$

where  $C_i^I$  is the total carbon emissions of the industry  $i$  of province  $I$ ,  $CE_{ik}^I$  is the carbon emitted by industry  $i$  of province  $I$  consuming energy  $k$ ,  $AD_{ik}^I$  represents the total amount of energy  $k$  consumed by the industry  $i$  of province  $I$ ,  $NCV_k^I$  represents the average low calorific value of energy  $k$ ,  $CC_k^I$  is the carbon emissions per unit of heat, and  $O_{ik}^I$  is the oxidation rate.

Then the total carbon emissions caused by final consumption is

$$CE = \hat{E} (I - A)^{-1} Y, \quad (6)$$

where  $\hat{E}$  is the carbon emissions intensity coefficient matrix ( $E_i^I = C_i^I / X_i^I$ ):

$$\hat{E} = \begin{pmatrix} E_1^1 & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \vdots & \vdots & 0 \\ \vdots & \cdots & E_i^I & \vdots & \vdots \\ \vdots & \cdots & \cdots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & E_m^n \end{pmatrix}. \quad (7)$$

From the producers' perspective, the carbon emissions  $CE_p^I$  represent the CO<sub>2</sub> emissions produced by the production activity in province  $I$ :

$$CE_p^I = \sum_{K=1}^m \hat{E}^I (I - A)^{-1} Y^K. \quad (8)$$

From the consumers' perspective, the carbon emissions  $CE_c^I$  represents the CO<sub>2</sub> emissions from production activities in all provinces caused by province  $I$ 's demand for products and services:

$$CE_c^I = \sum_{K=1}^m \hat{E}^K (I - A)^{-1} Y^I. \quad (9)$$

Interprovincial carbon emissions transfer originates from interprovincial trade. That is, in order to meet the needs of other provinces and cities for products and services, one province or city needs to conduct a production activity. Then the amount of carbon emissions transfer  $CT^{I,J}$  is represented by the carbon emissions generated during the production process. It means that the carbon emissions of various industries in province  $I$  caused by the final use by province  $J$  is the carbon emissions transfer from province  $J$  to province  $I$ :

$$CT^{I,J} = \hat{E}^I (I - A)^{-1} Y^J. \quad (10)$$

*3.2. Carbon Emissions Results from Different Perspectives.* This study adopted the interregional input-output tables of 42 sectors in 31 provinces, municipalities, and autonomous regions in China in 2012, 2015, and 2017. According to the method of Chen et al. [12], the 42 sectors in the input-output tables were merged into 16 sectors. The energy consumption by each sector, each region in each year was from the official data released by CEADs (<https://www.ceads.net>). The results are shown in Table 1.

In 2012, 2015, and 2017, the highest total carbon emissions from the consumers' perspective were coastal provinces with large populations and high consumption levels, such as Shandong (566 million tons per year) and Guangdong (528 million tons per year). The lower ones were in regions with relatively small populations and low consumption levels including Tibet (15 million tons per year) and Qinghai (41 million tons per year). The primary reason is the large differences in the total population and consumption levels between regions. A large population and a high consumption level make the total final use large in that region, leading to high-carbon emissions from the consumers' perspective. It is the opposite way in sparsely populated areas with low consumption levels.

With respect to the time trend, benefit from the increase in consumption levels by economic development during the period, the total carbon emissions from the consumers' perspective in populous regions such as Guangdong, Jiangsu, and Henan were all rising. While regions such as Beijing, Shanghai, and Tianjin which already have high consumption levels demonstrated an overall decreasing pattern from the perspective of consumers. It indicated that economic development in areas with high consumption levels has brought about a low-carbon consumption structure.

From the perspective of the producers, the high total carbon emissions were in regions like Shandong (645 million tons per year), Hebei (556 million tons per year), and Shanxi (379 million tons per year). The industrial structure in these regions preferred high polluting industries such as steel and energy. The advantageous industries in Shandong include three high-polluting and high-energy-consuming industries, energy processing (petroleum and coal), nonferrous extractive metallurgy, and chemical raw material production, which resulted in much higher carbon emissions than other regions. The lower total carbon emissions from the perspective of producers were regions with advantages in the tertiary industry, such as Beijing (58 million tons per year), Hainan (35 million tons per year), and Tibet (0.05 million tons per year).

Regarding the time trend, the total carbon emissions were relatively stable from the perspective of producers, different from those of the perspective of consumers. Except for Inner Mongolia and Xinjiang whose changes were relatively big, the changes in total carbon emissions were small in other provinces/cities. The primary reason is that their

industrial structure is stable compared to the consumption level; therefore, the changes in total carbon emissions caused by industrial structure were small.

*3.3. Total Net Carbon Emissions Transfer.* According to the principle of "the consumers assume responsibility," the net carbon emissions transfer was calculated by subtracting the producers' total carbon emissions from the consumers' total carbon emissions, that is, equation (9) minus equation (8).

The regions with high annual net carbon emissions transfer-out in 2012, 2015 and 2017 were provinces/cities with large populations and high consumption levels such as Guangdong, Zhejiang, and Beijing. Their industrial structure could not support their own demand, so they purchased products from other provinces/cities and transferred the CO<sub>2</sub> that should be emitted in region for producing these products to other regions. The regions with high annual net carbon emissions transfer-in were regions with an industry structure preferring energy and heavy-industry, such as Inner Mongolia, Hebei, and Shanxi. On the one hand, the energy industry, as an important part of production and life, participates in the production process of other industries in the region. On the other hand, the region's own demand is far from releasing its production capacity. While a large number of products were sold by these regions, the CO<sub>2</sub> that should be emitted by other provinces/cities for producing these products was transferred in. With respect to the time trend, provinces/cities with negative annual net carbon emissions transfer have greater fluctuations compared to those with positive annual net carbon transfer. Especially the fluctuations in Chongqing, Yunnan, Shanghai, and Henan were big. While Shanghai and Henan even changed the direction of net carbon transfer.

By equation (10), we can get the interprovincial carbon emission transfer matrix, but it is difficult to display due to the large amount of data. To visualize the carbon emissions transfer relationship clearly between provinces/cities, the amount of carbon transfer from one province/city to the other provinces/cities was sorted by the proportion of the total transfer amount, and the top 60% was taken to generate the interprovincial carbon emissions transfer relationship matrix. Then the social network was drawn with arrows pointed the direction of CO<sub>2</sub> transfer (see Figure 1).

In general, the interprovincial carbon emissions transfer social network exhibited a star-like shape. The five provinces of Inner Mongolia (IM), Hebei (HB), Shanxi (SX), Liaoning (LN), and Shandong (SD) with high net carbon transfer-in not only received most transfer-out from Guangdong (GD), Zhejiang (ZJ), Beijing (BJ), Chongqing (CQ), Shanghai (SH), and Tianjin (TJ), the regions with high net carbon transfer-out but also served as the main transfer-out places of carbon emissions from other provinces/cities. They worked as the center of the interprovincial carbon emissions transfer network and were also the dominant regions from the perspective of producers. From the perspective of coordinated governance of regional carbon emissions, the impact of the emissions reduction policy on the central provinces could quickly spread through the transfer network, conducive to timely adjustment of policy based on actual effects.



TABLE 1: Carbon emissions results from different perspectives.

Year	Consumers' perspective			Producers' perspective		
	2012	2015	2017	2012	2015	2017
Beijing	209.0	170.9	142.7	65.5	59.6	49.4
Tianjin	202.0	199.7	117.4	112.5	112.5	104.1
Hebei	351.9	369.2	476.1	536.6	556.0	576.5
Shanxi	227.8	209.6	241.6	378.1	355.2	406.4
Inner Mongolia	291.0	219.1	263.9	455.9	463.6	549.2
Liaoning	302.5	304.7	240.1	356.3	371.0	371.4
Jilin	223.0	209.4	175.7	205.3	185.2	179.6
Heilongjiang	261.8	223.8	239.3	222.6	228.6	232.5
Shanghai	277.4	161.7	122.9	123.0	115.1	132.3
Jiangsu	409.9	431.7	517.5	436.5	504.5	515.9
Zhejiang	320.1	420.1	435.8	252.7	264.0	258.3
Anhui	185.0	257.9	246.7	258.1	289.6	310.2
Fujian	191.2	148.2	135.4	165.8	169.7	178.9
Jiangxi	128.0	184.3	196.9	128.7	166.7	182.0
Shandong	566.4	602.0	532.1	640.0	662.0	635.6
Henan	356.9	426.0	517.8	419.8	435.1	407.7
Hubei	349.5	379.0	316.4	310.7	269.5	277.9
Hunan	241.4	291.7	351.7	237.0	239.5	265.4
Guangdong	559.9	438.0	586.0	298.6	294.5	345.0
Guangxi	203.2	202.3	175.7	173.4	171.6	186.0
Hainan	50.1	53.9	36.0	32.4	36.3	37.1
Chongqing	159.0	291.6	197.1	138.8	131.4	124.4
Sichuan	233.1	269.7	289.6	258.7	272.8	263.8
Guizhou	131.0	161.3	199.5	182.4	200.1	209.5
Yunnan	200.2	244.7	278.1	176.5	150.4	169.0
Tibet	12.4	8.5	24.6	4.0	4.4	6.6
Shaanxi	220.6	238.1	270.1	191.8	227.2	217.6
Gansu	90.7	105.9	96.5	121.5	132.7	126.1
Qinhai	34.6	36.8	54.5	34.0	42.1	46.9
Ningxia	50.4	56.0	99.4	108.9	119.1	145.6
Xinjiang	181.3	191.5	273.7	195.2	277.1	339.7

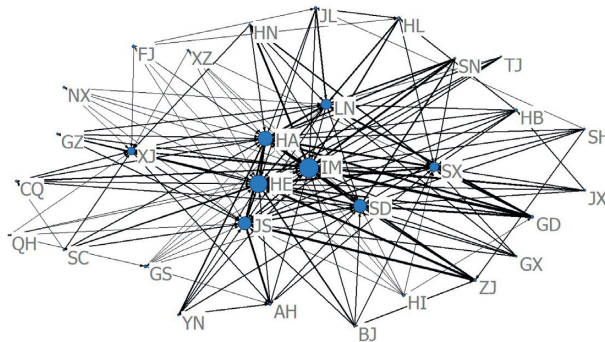


FIGURE 1: Interprovincial carbon emissions transfer social network (2017).

Although the net carbon transfer-in Henan was not high in 2015, and the net carbon emissions were transferred-out in 2017, it is still close to the center of the carbon emissions transfer network. It is because that the carbon emissions transferred to the “sub-center” Henan were far less than the amount transferred to the “central” provinces, and the carbon emissions transferred from Henan to other provinces were high too. Similarly, Jiangsu had a net carbon transfer-out of only 0.02 million tons in 2017, whereas it was still close to the center of the network. The main reason is also that provinces such as Inner Mongolia, Shanxi, and Liaoning

received most of the carbon emissions transfer, while carbon emissions transferred from Jiangsu’s to other provinces was high too.

#### 4. Decomposition of Factors Affecting Net Carbon Emissions Transfer

To support the construction of a coordinated governance mechanism for regional carbon emissions, it is not only necessary to grasp the current interprovincial carbon emissions transfer and carbon transfer network but also

need to fully understand the continuous and systemic influencing factors of net carbon emissions transfer. Thus, this article adopted the LMDI method to decompose the influencing factors of the net carbon emissions transfer at the provincial level.

**4.1. Establishment of the LMDI Model.** In order to explore the key factors affecting the net carbon emission transfer in each province, the contribution of various factors to the carbon emission transfer was evaluated. Since this study focused on the impact of carbon emissions intensity, output structure, demand-output ratio, and total demand on carbon emissions, the following equation is used for analysis:

$$C_{ex}^I = \sum_{j=1}^m \sum_{j=1}^n \left( \frac{C_j^{(I,J)}}{X_j^{(I,J)}} \right) \left( \frac{X_j^{(I,J)}}{X^{(I,I)}} \right) \left( \frac{X^{(I,I)}}{Y_T^I} \right) Y_T^I, \quad (11)$$

where  $E_{ex}^{(I,J)j} = C_j^{(I,J)}/X_j^{(I,J)}$ , representing the carbon emissions transferred per unit of output from province  $I$  to industry  $j$  of province  $J$ ;  $ST_{ex}^{(I,J)j} = X_j^{(I,J)}/X^{(I,I)}$ , representing the industry structure of production activities of province  $J$

to meet the needs of province  $I$ ;  $F_{ex}^{(I,J)j} = X^{(I,I)}/Y_T^I$ , representing the output of province  $J$  caused by per unit of final demand of province  $I$ ; and  $Y_{ex}^{(I,J)} = Y_T^I$ , representing the final demand scale of province  $I$ .

$$C_{im}^I = \sum_{j=1}^m \sum_{j=1}^n \frac{C_j^{(I,I)}}{X_j^{(I,I)}} \frac{X_j^{(I,I)}}{X^{(I,I)}} \frac{X^{(I,I)}}{Y_T^I} Y_T^I, \quad (12)$$

where  $E_{im}^{(I,I)j} = C_j^{(I,I)}/X_j^{(I,I)}$ , representing the carbon emissions transferred per unit of output from industry  $j$  of province  $J$  to province  $I$ ;  $ST_{im}^{(I,I)j} = X_j^{(I,I)}/X^{(I,I)}$ , representing the industry structure of production activities of province  $I$  to meet the needs of province  $J$ ;  $F_{im}^{(I,I)j} = X^{(I,I)}/Y_T^I$ , representing the output of province  $I$  caused by per unit of final demand of province  $J$ ; and  $Y_{im}^{(I,I)} = Y_T^I$ , representing the final demand scale of province  $J$ .

This article employed the LMDI method to decompose the factors that affect the provincial net carbon emissions into technical effect, structural effect, input-output effect, and scale effect:

$$\begin{aligned} \Delta C^I &= C_{ex}^I - C_{im}^I \\ &= \Delta C_E^I + \Delta C_{ST}^I + \Delta C_F^I + \Delta C_Y^I, \\ \Delta C_E^I &= \sum_{j=1}^m \sum_{j=1}^n W^{(I,J)j} \ln \frac{E_{ex}^{(I,J)j}}{E_{im}^{(I,I)j}} \dots \dots \text{technical effect}, \\ \Delta C_{ST}^I &= \sum_{j=1}^m \sum_{j=1}^n W^{(I,J)j} \ln \frac{ST_{ex}^{(I,J)j}}{ST_{im}^{(I,I)j}} \dots \dots \text{structural effect}, \\ \Delta C_F^I &= \sum_{j=1}^m \sum_{j=1}^n W^{(I,J)j} \ln \frac{F_{ex}^{(I,J)j}}{F_{im}^{(I,I)j}} \dots \dots \text{input - output effect}, \\ \Delta C_Y^I &= \sum_{j=1}^m \sum_{j=1}^n W^{(I,J)j} \ln \frac{Y_{ex}^{(I,J)}}{Y_{im}^{(I,I)}} \dots \dots \text{scale effect}, \end{aligned} \quad (13)$$

where  $W^{(I,J)j} = \begin{cases} (C_j^{(I,J)} - C_j^{(J,I)}) / \ln(C_j^{(I,J)} - C_j^{(J,I)}), & C_j^{(I,J)} \neq C_j^{(J,I)} \\ C_j^{(I,J)}, & C_j^{(I,J)} = C_j^{(J,I)} \end{cases}$ .

The technology effect reflects the differences in production technology between provinces in interprovincial trade. The structural effect reflects the contribution of the difference in the final demand structure of each province to the transfer of net carbon emissions. The input-output effect reflects the degree of correlation between the various industries in each province in the production process. The scale effect reflects the contribution of differences in the scale of final demand across provinces to the transfer of net carbon emissions implied by trade.

**4.2. Analysis of LMDI Decomposition Results.** This paper decomposed the annual net carbon transfer into technical effect, structural effect, input-output effect, and scale effect

via LMDI method. A positive contribution of an effect indicated that this effect made the regional carbon transfer-out greater than carbon transfer-in, and a negative contribution value indicated a larger carbon transfer-in than carbon transfer-out. The decomposition results are shown in Table 2. Due to the large amount of data, only data of 2017 are displayed.

The provinces/cities with a positive technical effect were concentrated in the economy developed regions such as Beijing, Guangdong, and Zhejiang, as well as the regions whose economic development did not rely on high-polluting and high-emission industries, such as Jiangxi and Yunnan. The main reasons might be that, the regions such as Beijing and Guangdong, have advanced carbon emission technology with developed economy compared to their trading regions, resulting smaller amount of carbon emissions of the

TABLE 2: LMDI decomposition results in 2017.

Province	Total effect	Total effect	Total effect	Technical effect	Structural effect	Input-output effect	Scale effect
Year	2012	2015	2017			2017	
Beijing	1.44	1.11	0.93	1.30	--0.17	-0.15	-0.05
Tianjin	0.90	0.87	0.13	0.35	-0.03	0.10	-0.29
Hebei	-1.85	-1.87	-1.00	-0.63	0.15	-0.80	0.27
Shanxi	-1.50	-1.46	-1.65	-0.79	-0.92	0.85	-0.78
Inner Mongolia	-1.65	-2.44	-2.85	-1.55	-1.34	1.09	-1.05
Liaoning	-0.54	-0.66	-1.31	-0.58	-0.55	0.14	-0.32
Jilin	0.18	0.24	-0.04	-0.17	0.39	0.35	-0.61
Heilongjiang	0.39	-0.05	0.07	-0.68	0.18	1.00	-0.43
Shanghai	1.54	0.47	-0.09	0.22	0.37	-0.45	-0.24
Jiangsu	-0.27	-0.73	0.02	0.31	0.76	-3.05	2.00
Zhejiang	0.67	1.56	1.78	0.96	0.21	-0.34	0.93
Anhui	-0.73	-0.32	-0.64	0.04	0.20	-0.63	-0.24
Fujian	0.25	-0.22	-0.43	0.23	-0.33	-0.33	-0.01
Jiangxi	-0.01	0.18	0.15	0.25	0.19	0.13	-0.41
Shandong	-0.74	-0.60	-1.04	0.04	-0.44	-1.99	1.35
Henan	-0.63	-0.09	1.10	0.70	0.62	-1.76	1.55
Hubei	0.39	1.09	0.38	0.00	0.02	0.17	0.20
Hunan	0.04	0.52	0.86	0.15	0.59	-0.09	0.22
Guangdong	2.61	1.43	2.41	1.28	1.31	-2.10	1.92
Guangxi	0.30	0.31	-0.10	-0.11	-0.03	0.34	-0.30
Hainan	0.18	0.18	-0.01	-0.02	-0.01	0.44	-0.42
Chongqing	0.20	1.60	0.73	0.13	0.70	0.28	-0.39
Sichuan	-0.26	-0.03	0.26	0.23	0.04	-0.20	0.19
Guizhou	-0.51	-0.39	-0.10	-0.07	-0.34	0.86	-0.55
Yunnan	0.24	0.94	1.09	0.06	0.04	1.07	-0.08
Tibet	0.08	0.04	0.18	0.09	0.05	0.11	-0.07
Shaanxi	0.29	0.11	0.53	-0.16	0.44	0.57	-0.32
Gansu	-0.31	-0.27	-0.30	-0.10	-0.31	0.66	-0.54
Qinghai	0.01	-0.05	0.08	0.00	-0.06	0.30	-0.18
Ningxia	-0.58	-0.63	-0.46	-0.42	-0.76	1.52	-0.80
Xinjiang	-0.14	-0.86	-0.66	-1.05	-0.93	1.89	-0.57

corresponding equivalent commodities. Therefore, the carbon transfer-in caused by production was less than the carbon transfer-out caused by consumption. Provinces and cities that do not rely on high-polluting and high-emission industry for economic development such as Jiangxi and Yunnan had advantages in carbon emissions intensity when trading with regions with a large proportion of heavy industry, leading to net carbon emissions transfer-out. On this basis, during the coordinated governance of regional carbon emissions, we could encourage technology exchanges between regions with positive technical effects and regions with negative technical effect to speed up the spread of current low-carbon production technologies. In addition, we could strengthen support for technological innovation in regional with negative technical effect to guide the production technology toward low-carbon style.

The regions with positive structural effect included 16 provinces and cities, such as Jiangsu, Chongqing, Shanghai, and Tianjin. The main reason is that the provinces and cities, such as Jiangsu, Chongqing, and Shanghai, when trading with other regions, gained output in industries with low-carbon emissions such as transportation and storage and construction in their own regions. While the provinces such as Inner Mongolia, Shanxi, and Ningxia, gained output in industries with high-carbon emissions such as coal mining

and dressing, metal smelting and pressing, and power and heat supply. Therefore, the regions with positive structural effect would have a smaller carbon emissions transfer-in from production than the carbon emissions transfer-out from consumption, thus a net carbon transfer-out. Whereas the regions with negative structural effect would have a larger carbon transfer-in than transfer-out, thus a new carbon transfer-in. From the perspective of collaborative governance, it is necessary to fully consider the differences in resource endowments between provinces and cities, to guide the industrial structure of each region to give full play to its own advantages, and to rationally lead the transfer of high-polluting and high-emission industries between regions.

The regions with positive input-output effect were mainly distributed in areas with relatively immature industrial chains and weak economies such as the northwest and southwest. While regions with negative input-output effect were mainly distributed in the eastern and northern coastal areas, where the industrial chains were more mature and the input-output relationship between industries was more complex. When meeting the needs of other provinces, the total output increase of all industries brought by the equivalent demand from other provinces will be higher than that of regions with weaker inter-industry linkages. It is likely to cause larger carbon transfer-in from production

than carbon transfer-out from consumption, thus the net carbon transfer-in. The input-output effect was the dominant factor of net carbon transfer. Therefore, in the overall governance process, it is necessary to rationally consider the economic benefits of the increased production of various industries brought about by the interprovincial trade demand and the carbon emissions transfer that comes along, so as to improve the interprovincial industrial chain. Optimizing the industrial structure and leading the input-output relationship between industries toward low-carbon can be an effective practical direction to promote the coordinated governance of regional carbon emissions.

The regions with positive scale effect were mainly distributed in developed coastal provinces. They have large populations and high consumption levels, and the corresponding final consumption scales were larger than other provinces. Therefore, the scale of consumption would be greater than the scale of production in interprovincial trade, and the carbon emissions transfer-out from consumption would be larger than transfer-in from production, resulting in a net carbon transfer-out. From the perspective of coordinated governance of regional carbon emissions, it is necessary to build a reasonable horizontal inter-provincial carbon compensation mechanism, so that high-consumption regions could assume the carbon emission responsibility of consumed goods, and build a bridge of coordinated carbon governance between carbon transfer-out and transfer-in provinces and cities.

In terms of time trend, among the provinces with large changes in carbon emissions transfers in 2012, 2015, and 2017, the changes in total effect in Inner Mongolia, Liaoning, Jilin, Heilongjiang, Hunan, and Zhejiang were mainly due to the same direction changes of scale effect. The decline in total effect in Fujian was primarily due to the reduction in input-output effect. While for provinces and cities with drastic changes in the amount of carbon emissions transfer, the changes in total effect in Shanghai, Jiangsu, and Henan were mainly caused by the same direction changes of scale effect. It can be seen that the scale effect was the dominant factor causing changes in carbon emissions transfer. Presumably, it is because that the final consumption scale responded more quickly to changes in environmental factors such as consumers' preferences and industry policies than the industrial structure, technological level, and industry input-output relationships.

## 5. Conclusion and Suggestion

This study constructed an interregional input-output model for 31 provinces and cities in China to calculate and analyze the carbon emissions transfer between provinces and cities, carbon emissions from the perspective of producers and from the perspective of consumers. Furthermore, this study conducted influencing factor decomposition for net carbon emissions transfer using the LMDI method. The following conclusions are drawn:

- (1) High-carbon emission areas from the consumers' perspective are mainly distributed in populous provinces and cities such as Shandong, Guangdong, Jiangsu, and Henan, and the emissions are rising. While in regions with high consumption levels such as Beijing, Shanghai, and Tianjin, the carbon emissions are declining from the consumers' perspective. High-carbon emission areas from the producers' perspective are mainly distributed in regions with industrial structure preference toward high-polluting industries like steel and energy, including Shandong, Hebei, and Shanxi. And the carbon emissions the producers' perspective fluctuated little over time and are relatively stable.
- (2) The interprovincial carbon emissions transfer network is star-like. The five provinces, Inner Mongolia, Hebei, Shanxi, Liaoning, and Shandong, with high net carbon transfer-in are the primary transfer-out places of carbon emissions from other provinces, and they are located in the center of the carbon emissions transfer network.
- (3) The input-output effect is the dominant one that affects the net carbon transfer, indicating that the net carbon transfer is primarily affected by the completeness degree of the industrial chain. In the eastern and northern coastal areas, the industrial chain is relatively complete, and the input-output relationship between industries is more complicated. Their increases in total industrial output caused by the same demand are higher than that of other regions. Therefore, the carbon transfer-in from production is greater than the carbon transfer-out from consumption, resulting a net carbon transfer-in. The scale effect is the primary factor determining the carbon emissions transfer from 2012 to 2017. The reason for that is that compared with the industrial structure, technological level, and industry input-output relationships, the final consumption scale responds more quickly to changes in environmental factors including consumers' preferences and industry policies.

Based on the above-mentioned empirical results and research conclusions, the following suggestions are made. First, we should strengthen support for low-carbon industries and low-carbon technologies. On the one hand, to provide more support for low-carbon industries such as clean energy to improve energy efficiency and expedite the low-carbon shift in industrial structure. On the other hand, to encourage low-carbon technological innovation and lead high-polluting and high-emission industries to reduce pollution and emissions. At the same time, to promote low-carbon consumption and low-carbon lifestyle and guide the low-carbon shift in consumption concept. We should take full advantage of the relationship between the center regions with other regions in the interprovincial carbon emissions

transfer network, prioritize the industrial upgrading and transformation of the metal and nonmetal processing and manufacturing industry in center regions, and eliminate outdated production capacity with high pollution and high emissions. Then, we should improve the responsibility mechanism for interprovincial carbon transfer. We should fully review the differences in economic development between provinces, comprehensively consider the producers' principle and consumers' principle, and rationally allocate carbon reduction responsibility in major high-carbon production regions and major high-carbon consumption regions. Last but not least, we should think about the differences in resource endowments of various provinces and cities, improve the industrial chain, and promote the low-carbon shift of input-output structure of inter-provincial industries.

### Data Availability

Energy consumption by each sector and each region in each year was taken from the official data released by CEADs (<https://www.ceads.net>).

### Conflicts of Interest

The authors declare no conflicts of interest.

### Authors' Contributions

Y.Q. designed the model and the computational frame work and analyzed the data. G.P. wrote the manuscript. Lei conducted empirical research. Lu collected the relevant literates. Y.P. collected the data. All authors discussed the results and contributed to the final manuscript. All authors have read and agreed to the published version of the manuscript.

### Acknowledgments

This research was supported by the National Natural Science Foundations of China (71763010 and 71803038) and Student Research Project of Jiangxi University of Finance and Economics (20210918091940252).

### References

- [1] J. Munksgaard and K. A. Pedersen, "CO<sub>2</sub> accounts for open economies: producer or consumer responsibility?" *Energy Policy*, vol. 29, no. 4, pp. 327–334, 2001.
- [2] Z. Zhong, L. Jiang, and P. Zhou, "Transnational transfer of carbon emissions embodied in trade: characteristics and determinants from a spatial perspective," *Energy*, vol. 147, pp. 858–875, 2018.
- [3] W. W. Leontief, "Quantitative input and output relations in the economic systems of the United States," *The Review of Economics and Statistics*, vol. 18, no. 3, pp. 105–125, 1936.
- [4] W. Leontief, "Environmental repercussions and the economic structure: an input-output approach," *The Review of Economics and Statistics*, vol. 52, no. 3, pp. 262–271, 1970.
- [5] J. e. Guo, Z. Zhang, and L. Meng, "China's provincial CO<sub>2</sub> emissions embodied in international and interprovincial trade," *Energy Policy*, vol. 42, pp. 486–497, 2012.
- [6] L. Yao, J. R. Liu, R. S. Wang, and K. Yin, "Carbon footprint accounting of regional household consumption in China through multi-regional input-output model," *Acta Scientiae Circumstantiae*, vol. 33, no. 7, pp. 2050–2058, 2013.
- [7] B. Zhang, H. Qiao, Z. M. Chen, and B. Chen, "Growth in embodied energy transfers via China's domestic trade: e," *Applied Energy*, vol. 184, pp. 1093–1105, 2016.
- [8] H. G. Liu, W. D. Liu, Z. P. Tang, and X. M. Fan, "The effect analysis of regional industry structure adjustment for CO<sub>2</sub> emission reduction in China: on the base of inter-regional input-output method," *Areal Research and Development*, vol. 29, no. 3, pp. 129–135, 2010.
- [9] B. Su and B. W. Ang, "Input-output analysis of CO<sub>2</sub> emissions embodied in trade: a multi-region model for China," *Applied Energy*, vol. 114, pp. 377–384, 2014.
- [10] B. Su and B. W. Ang, "Multiplicative structural decomposition analysis of aggregate embodied energy and emission intensities," *Energy Economics*, vol. 65, pp. 137–147, 2017.
- [11] C. Sun, D. Ding, and M. Yang, "Estimating the complete CO<sub>2</sub> emissions and the carbon intensity in India: from the carbon transfer perspective," *Energy Policy*, vol. 109, no. 10, pp. 418–427, 2017.
- [12] H. Chen, J. Wen, J. Pang, Z. Chen, and Y. S. Wei, "Research on the carbon transfer and carbon equity at provincial level of China based on MRIO model of 31 provinces," *China Environmental Science*, vol. 40, no. 12, pp. 5540–5550, 2020.
- [13] W. Wang and Y. Hu, "The measurement and influencing factors of carbon transfers embodied in inter-provincial trade in China," *Journal of Cleaner Production*, vol. 270, no. 1, Article ID 122460, 2020.
- [14] M. Han, Q. Yao, J. Lao, Z. Tang, and W. Liu, "China's intra- and inter-national carbon emission transfers by province: a nested network perspective," *Science China Earth Sciences*, vol. 63, no. 6, pp. 852–864, 2020.
- [15] X. E. Wang, S. H. Zhao, X. Y. Liu, H. Y. Duan, and J. N. Song, "Carbon neutrality oriented provincial consumption-based carbon emission reduction models: based on the multi-regional input-output model," *Ecological Economy*, vol. 37, no. 5, pp. 43–50, 2021.
- [16] H. G. Liu, Z. M. Zhang, and J. Guo, "Carbon emissions embodied in added-value chains in China," *Management Review*, vol. 33, no. 9, pp. 58–64, 2021.
- [17] Z. Q. Qian and L. K. Yang, "The impact of East Asia vertical specialization on China's embodied carbon emissions: an inter-temporal MRIO-SDA analysis," *Resources Science*, vol. 38, no. 9, pp. 1801–1809, 2016.
- [18] H. Jiang, "Implied carbon in trade between BRIC countries based on input-output modeling and structural decomposition," *Resources Science*, vol. 38, no. 12, pp. 2326–2337, 2016.
- [19] C. X. Guo, "Decomposition of China's carbon emissions: based on LMDI method," *China Population, Resources and Environment*, vol. 20, no. 12, pp. 4–9, 2010.
- [20] Y. S. Du and H. H. Sun, "On the growth factors of CO<sub>2</sub> embodied in China's export: an analysis based on LMDI," *World Economy Studies*, vol. 31, no. 11, pp. 44–49+88, 2012.
- [21] H. P. Huang, M. T. Yi, and J. W. Cao, "Trade embodied carbon emissions and its effects of China: take the Yangtze River Economic Belt as an example," *Economic Geography*, vol. 41, no. 3, pp. 49–57, 2021.

## *Retraction*

# **Retracted: Spare Parts Stocking Decision Strategy and Service Logistics Cost Optimization of Two-Echelon Service Logistics System considering Multifailure Mode**

### **Computational Intelligence and Neuroscience**

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### **References**

- [1] R. Wang and G. Chen, "Spare Parts Stocking Decision Strategy and Service Logistics Cost Optimization of Two-Echelon Service Logistics System considering Multifailure Mode," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 7607985, 12 pages, 2022.



## Research Article

# Spare Parts Stocking Decision Strategy and Service Logistics Cost Optimization of Two-Echelon Service Logistics System considering Multifailure Mode

Ruiqi Wang  and Guangyu Chen 

*School of Management and Economics, University of Electronic Science and Technology of China (UESTC), Chengdu, China*

Correspondence should be addressed to Guangyu Chen; 1211309064@qq.com

Received 18 January 2022; Accepted 4 February 2022; Published 8 March 2022

Academic Editor: Daqing Gong

Copyright © 2022 Ruiqi Wang and Guangyu Chen. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

For the repair level and spare parts stocking decision problems, generally METRIC type methods and level of repair analysis (LORA) are used separately. In practical engineering, the repair level of large-scale systems is usually judged according to the failure modes. The method of judging the repair level by the maintenance success rate is no longer applicable. In the case of multiple failure modes of a large-scale system, considering the requirements of system availability, we build a spare parts stocking decision and service logistics cost optimization model of a two-echelon service logistics system. Aiming at the spare parts stocking allocation problem caused by multiple failure modes, we improve the iterative greedy heuristic algorithm to find the global optimal stocking allocation strategies. Finally, through the analysis of typical examples, the correctness and effectiveness of the model and algorithm are verified. The impact of multifailure mode spare parts stocking allocation strategies on availability and service logistics cost is analyzed. The research results are helpful to simplify the support engineering design process of system engineers and have certain theoretical and application value.

## 1. Introduction

Large-scale system can be decoupled or decomposed into multiple interrelated subsystems or modules. Each subsystem contains a variety of components, which are controlled by the same constraints or objectives to complete specific tasks. Therefore, any component's failure will cause system failure [1, 2]. The large number of components means that a reasonable service logistics strategy should be developed to meet the availability of spare parts and the economy of the service logistics system [3].

In terms of economy, due to the complex operational environment of large-scale system, the establishment of maintenance center for unified resource allocation and repair failure units could save a lot of costs compared to on-site maintenance [4]. In terms of availability, on-site maintenance would save the transportation waiting time of failed units and improve system availability [5]. Therefore, the aim

of service logistics strategy is to balance the relationship between maintenance level, spare parts stocking allocation, and service logistics cost under the constraint of system availability.

The METRIC model is the most commonly used in service logistics strategy. The original METRIC model used "echelon" to represent the levels of different types of warehouses [6]. Slay [7] developed the VAR-METRIC model, which assumes that the average number of components in repair is equal to the variance of the negative binomial distribution. Graves [8] explored the same distribution, elaborating both an approximate and an exact method for the multiechelon single-indenture case. On this basis, literature [9] provides an accurate method to evaluate multiechelon service logistics systems. Therefore, the VAR-METRIC model is the most common method used by scholars to solve spare parts stocking allocation problems.

The level of repair analysis (LORA) is an input decision parameter of the VAR-METRIC model to solve the spare parts stocking allocation, and it is used to determine the location of spare parts for repair [10]. Falkner [11] was the first scholar to mention joint optimization of spare parts stocking decision and repair in his paper. Performing the LORA first and then the spare parts stocking decision analysis, the sequential approach, may lead to a solution that is not optimal. Basten [12] proposed an integrated algorithm to find optimal solutions for two-echelon problems. He uses efficient points on the curve of costs versus expected number of backorders to represent the optimal solution. On this basis, he proposed an iterative algorithm to solve the joint problem of LORA and spare parts stocking decision in 2015. The basic idea is to solve the LORA decision first, then use the VARI-METRIC model to solve the spare parts stocking decision problem, and last use the results of VARI-METRIC to add an estimate of the holding costs to the LORA inputs and start a second iteration [10]. However, all the above studies assume that component failure is a single failure mode.

Large-scale systems have a large number of components, leading to a large number of complex failure mechanisms. That is, different failure modes have different degrees of influence on the system [13]. To define or distinguish these failure modes, many scholars will choose failure mode analysis according to failure mechanism, failure impact, failure rate, or other parameters [14]. Failure mode analysis is an important work item of system availability analysis, and it is the basis for maintainability analysis, safety analysis, testability analysis, and supportability analysis [15]. Therefore, before formulating LRU maintenance level and spare parts stocking decision strategies, failure modes analysis of key component should be done first.

The introduction of multiple failure modes would change the joint optimization model of METRIC and LORA. Different echelon warehouses need different types of spare parts. Therefore, a new two-echelon service logistics cost optimization model for large-scale system is built in this paper. We improve an iterative greedy heuristic algorithm based on literature [10]. Finally, the correctness and effectiveness of the model and algorithm are verified by a practical engineering case. The key influencing factors of spare parts stocking allocation decisions and system availability are revealed. This research has certain theoretical and application value in the optimization of service logistics strategy.

## 2. Model Description

This section describes the characteristics of two-level support system for a large-scale operation system. According to the current research, we consider the influence of multiple failure modes on system operation results and the impact on maintenance strategy of support system. It provides a theoretical basis for the construction of the spare parts stocking decisions optimization model.

*2.1. System Description.* According to the hierarchical structure of Figure 1, this paper proposes a large-scale operation system with multiple failure modes. The system contains  $k$  identical individual subsystems, and each subsystem consists of  $i$  ( $i = 1, 2, 3, \dots$ ) series components.  $LRU_i$  is the key component of the subsystem. In case of failure, it can be repaired at the maintenance site of the support system.

The structure of large-scale system is complex, and LRU often has a variety of failure modes  $FM_i$ ,  $n$  ( $n = 1, 2, 3, \dots$ ) [16–18].

The key components of subsystems are various, including electronic, mechanical, and optical components. A single failure mode cannot meet the requirements of system availability analysis. At present, scholars generally use three classification methods: the impact of failure on components, the impact of failure on maintenance strategies, and the impact of failure on operating system. The first type divides the failure into degraded function failure, nonfunction failure, partial function failure, intermittent function failure, and unexpected function failure [19, 20]. The second type divides the failure into maintainable failure and non-maintainable failure [21]. The third type divides failures into hard failure and soft failure [22–24].

Considering that our optimization model is a service logistics model, we pay more attention to the impact of spare parts stocking decision strategies on system operation results. Therefore, the third type is selected as the classification method in this paper. Soft failure usually causes system production reduction or inferior performance, while the system keeps functioning; hard failure can cause system failure, for example, the service logistics model of an offshore wind farm in literature [5]. According to the difficulty of maintenance caused by the marine environment and the impact of turbine failure on the system, hard failure and soft failure classification methods are used to formulate maintenance strategies for different failure modes.

Specifically, we ascribe the sudden failure modes such as circuit break, short circuit, and impact to hard failure, which will lead to the instantaneous loss of system function. We ascribe wear, aging, fatigue, and other degradation failure modes to soft failure, which reduces the key performance of the system. According to the above analysis, the basic assumptions of large-scale operation system in this study are as follows:

- (1)  $LRU_i$  ( $i = 1, 2, 3, \dots$ ) are independent of each other, and the demand for spare parts follows the Poisson distribution.
- (2) System failure would be caused by one LRU at most. Multiple failures that occur at the same time are not considered.
- (3) When there is a shortage of spare parts in a failure maintenance site, the system stops, and the system availability decreases.
- (4) The importance of all LRUs is the same, and the priority is not considered for maintenance.

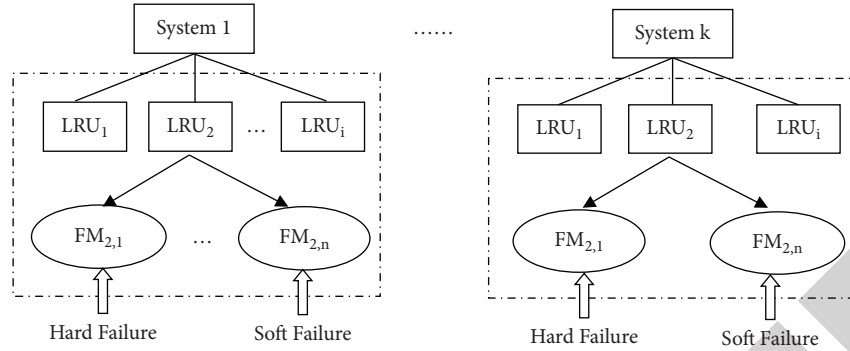


FIGURE 1: The large-scale operation system with multiple failure modes.

- (5) Failure diagnosis time and spare parts replacement time are ignored.
- (6) Each LRU has multiple failure modes  $FM_i$ ,  $n$  ( $n = 1, 2, 3, \dots$ ). The failure modes are mutually exclusive; that is, the simultaneous occurrence of multiple failure modes is not considered.
- (7) Failure modes are divided into hard failure and soft failure.

2.2. Two-Echelon Service Logistics System Description. When the operating system fails, the corresponding failed spare parts shall be replaced immediately, and the removed failed spare parts shall be sent to the maintenance site for repair [17]. In the level of repair analysis (LORA), the expensive resource cost usually leads to the repair work upstream of the maintenance network, while the expensive transportation cost leads to the downstream repair work close to the installation Base. Therefore, spare parts with cheap resources should generally be repaired locally, while spare parts requiring expensive resources should be repaired intensively [16].

Considering that the hard failure is mainly caused by overload, instantaneous impact, and electronic failure, it is assumed that the “Base” site of the service logistics system can completely repair this failure mode. However, the soft failure needs to be reprocessed and assembled. In such cases, maintenance activities are expensive to deploy to the complex operating environment on the Base site. Therefore, we assume that the soft failed spare parts are sent to a higher-grade maintenance site “Depot.” In this study, the purpose of considering multiple failure modes is to solve the problem of choosing the maintenance position of multiple spare parts types, which can make up for the spare parts allocation problem that LORA can only solve a single failure mode. Using failure mode characteristics to determine the maintenance level is more suitable for practical engineering.

The maintenance activities at different maintenance sites are different. We consider a two-echelon service logistics system. This system’s availability depends on the availability provided by any single maintenance site [17], and its structure is shown in Figure 2.

At each maintenance site of the service logistics system, maintenance resources such as maintenance tools,



FIGURE 2: The structure of two-echelon service logistics system.

maintenance personnel, and support equipment are configured to enable it to deal with different failure modes. The Base is usually set at the site of the operating system. In case of failure, if the spare part is in stocking in the Base, the spare part shall be extracted and replaced immediately to restore system availability. If the spare parts are out of stocking, an order will be issued to the Depot, and an  $(s - 1, s)$  replenishment strategy will be formulated [12, 17].

The maintenance of failed spare parts requires inspection to determine the failure mode and relevant maintenance activities: LRU with hard failure is repaired in Base, and LRU with soft failure is repaired in Depot. The Depot can be regarded as a warehouse with unlimited capacity. However, due to cost control, we still need to optimize the amount of LRUs stored in maintenance sites. When the failed spare part is repaired in the Depot, it will be sent to the source Base again [16].

According to the above analysis, the basic assumptions of the two-echelon service logistics system are as follows:

- (1) The maintenance site has the maintenance ability to deal with the corresponding failure mode; that is, discard decision is not considered
- (2) The hard failure shall be repaired in Base, and the soft failure can only be repaired in Depot
- (3) Base and Depot deal with different failures, their resource allocation and maintenance difficulties are different, and the repair time of failures is independent
- (4) The maintenance strategy selects perfect maintenance; that is, LRU is repaired as good as new
- (5) There are no lateral transshipments between locations at the same echelon level or emergency shipments from locations at a higher echelon level
- (6) Maintenance resources are nonconsumables and can be reused

The purpose of failure mode classification is to distinguish the types of spare parts. Under different failure modes, LRU needs different types of spare parts, different maintenance difficulties, and different maintenance positions. The optimization model in this study is also applicable to other failure mode classification methods. As long as the failure mode characteristics can be determined, the maintenance level problem can be solved according to the maintenance requirements of spare parts.

### 3. Two-Echelon Service Logistics Cost Optimal Modelling

The purpose of constructing this optimization model is to find the spare parts stocking allocation strategy that makes the service logistics cost optimal under the condition of meeting the requirements of system availability. We improve the LORA and METRIC joint model and design a new iterative algorithm. The model variables are shown in Table 1.

*3.1. Mathematical Model of Spare Parts Stocking Decision.* Generally, a variable is defined in the METRIC model to represent the maintenance capability of the maintenance site, representing the probability of successful repair by LRU<sub>*i*</sub> at *j*. This method is applicable to the case of single failure mode. In the case of multiple failure modes, the maintenance level should be determined in advance so that each maintenance point has sufficient capacity to deal with the current failure. This idea can not only clarify the resource allocation of different maintenance sites but also improve the efficiency of system availability management.

In this paper, the availability of the service logistics system is selected as the index to measure the effectiveness of the spare parts stocking decision [25]. The spare parts stocking holding cost optimization model of the two-echelon service logistics system is as follows:

$$\begin{cases} \min, & \sum \sum f c_{i,j} \cdot s_{i,j}, \\ s.t., & A(s_{i,j}) \geq A_{\min}. \end{cases} \quad (1)$$

Equation (1) indicates that the optimal spare parts stocking holding cost should meet the availability requirements  $A_{\min}$ .

When there are no LRU spare parts to be replaced in the warehouse of the maintenance site, the LRU will enter the shortage state, which will lead to the shutdown of the operating system and the decline of system availability. The relationship between availability and shortage of maintenance sites is as follows [26]:

$$A(s_{i,j}) = \left( 1 - \frac{\text{EBO}(s_{i,j})}{M_{i,j}} \right)^{M_{i,j}}, \quad (2)$$

where  $M_{i,j}$  is the assembly quantity of LRU<sub>*i*</sub> at *j*.

Since the failure of any LRU will cause system shutdown, it can be considered that the LRUs are in a series structure. The availability of the service logistics system  $A_{\text{sys}}$  is the product of the availability of all maintenance sites:

TABLE 1: Model variables.

$i$	Item number, LRU <sub><i>i</i></sub> , $i = 1, 2, 3, \dots, n$
$j$	Site number, Depot: $j = 0$ , Base: $j = 1, 2, 3, \dots, n$
$\lambda_{i,j}$	Single cycle demand mean value of item <i>i</i> at site <i>j</i>
$\lambda_{ih}$	Single cycle demand mean value of item <i>i</i> 's hard failure mode
$\lambda_{is}$	Single cycle demand mean value of item <i>i</i> 's soft failure mode
$T_{i,j}$	Repairing time of item <i>i</i> at site <i>j</i>
$O_{i,j}$	Order and ship time of item <i>i</i> from the Depot to the site <i>j</i>
$f c_i$	Price of item <i>i</i>
$s_{i,j}$	Level of stocking of item <i>i</i> at site <i>j</i>

$$A_{\text{sys}} = \prod_{j=1}^J \prod_{i=1}^I A(s_{i,j}) = \prod_{j=1}^J \prod_{i=1}^I \left( 1 - \frac{\text{EBO}(s_{i,j})}{M_{i,j}} \right)^{M_{i,j}}. \quad (3)$$

Take logarithms on the left and right sides of (3) to obtain

$$\begin{aligned} \ln A_{\text{sys}} &= M_{i,j} \cdot \ln \left( 1 - \frac{\text{EBO}(s_{i,j})}{M_{i,j}} \right) \approx M_{i,j} \cdot \left( -\frac{\text{EBO}(s_{i,j})}{M_{i,j}} \right) \\ &= -\text{EBO}(s_{i,j}). \end{aligned} \quad (4)$$

Equation (4) can be simplified by using the equivalent infinitesimal principle. Because of  $A_{\text{sys}} \in [0, 1]$  and  $\text{EBO}(s_{i,j}) \geq 0$ , finding the maximum value of system availability is equivalent to finding the minimum value of stocking shortage. Therefore, the solution of availability is transformed into the solution of stocking shortage and finally into the solution of the spare parts stocking decision of the service system [27].

EBO represents the LRUs' expected shortage when the service system stocking level is  $s_{i,j}$ . EBO can be expressed as

$$\text{EBO}(s_{i,j}) = \sum_{x=s_{i,j}+1}^{\infty} (x - s_{i,j}) p(x|N_{i,j}), \quad (5)$$

where  $N_{i,j}$  is the intermediate variable of the optimization model. It represents the average number of LRUs in maintenance status. After maintenance activities, it can be provided to the system again as new spare parts for storage. According to the Palm theorem, when the LRU demand obeys the Poisson process with mean value  $\lambda$  and the mean repair time is  $T$ , the steady-state probability distribution of  $N$  obeys the Poisson distribution with the mean value  $\lambda T$  [26]. That means

$$N = \lambda \cdot T. \quad (6)$$

Set the number of LRU<sub>*i*</sub> failures in a single cycle to  $\lambda_i$ . After inspection, the number of hard failures is set to  $\lambda_{ih}$ , and the number of soft failures is set to  $\lambda_{is}$ . Their relationship can be expressed as follows:

$$\lambda_i = \lambda_{ih} + \lambda_{is}. \quad (7)$$

In the case of multiple failure modes, the expected shortage of LRU<sub>*i*</sub> should also be divided into Depot expected shortage  $\text{EBO}(s_{i,0})$  and Base expected shortage  $\text{EBO}(s_{i,j})$ .



Since the failed spare parts in soft failure mode are repaired in Depot, the number of parts under repair in Depot is

$$N_{i,0} = \lambda_{is} \cdot T_{i,0}. \quad (8)$$

Bring (8) into (5), and LRU<sub>i</sub> expected shortage at the Depot is

$$EBO(s_{i,0}) = \sum_{x=s_{i,0}+1}^{\infty} (x - s_{i,0})P(x|\lambda_{is} \times T_{i,0}), \quad (9)$$

where  $EBO(s_{i,0})$  is a stocking-related variable. When the operating system fails, a replenishment request is sent to the Base for the first time. If the Base is short of spare parts, a request would be sent to the Depot. Due to the ordering and transportation time  $O_{i,j}$ , the number of spare parts under repair of Base  $j$  should consider transportation spare parts:

$$N_{i,j} = N_{i,j}^1 + N_{i,j}^2. \quad (10)$$

The number of parts under repair  $N_{i,j}^1$  in Base is

$$N_{i,j}^1 = \lambda_{ih,j} \cdot T_{i,j}. \quad (11)$$

The number of parts under transportation  $N_{i,j}^2$  should consider two situations:

- (1) When Depot has spare parts stocking, the number of spare parts under repair is the number sent to Depot:

$$N_{i,j}^{2'} = \lambda_{is,j} \cdot O_{i,j}. \quad (12)$$

- (2) When the Depot is short of spare parts, the spare parts should be sent back to Base after being repaired at Depot. The average number of spare parts waiting for maintenance is equal to the expected shortage:

$$N_{i,j}^{2''} = \lambda_{is,j} \cdot f_{i,j} \cdot \frac{EBO(s_{i,0}|\lambda_{is}T_{i,0})}{\lambda_{is}}, \quad (13)$$

where  $f_{i,j} = \lambda_{is,j}/\lambda_{is}$ , which indicates the proportion of LRU<sub>i</sub> sent by Base  $j$ . Based on the above analysis, the number of spare parts under repair in a single cycle can be expressed as

$$\begin{aligned} N_{i,j} &= N_{i,j}^1 + N_{i,j}^{2'} + N_{i,j}^{2''} \\ &= \lambda_{ih,j} \cdot T_{i,j} + \lambda_{is,j} \left[ O_{i,j} + f_{i,j} \cdot \frac{EBO(s_{i,0}|\lambda_{is}T_{i,0})}{\lambda_{is}} \right]. \end{aligned} \quad (14)$$

Therefore, the expected shortage of Base is a variable related to stocking level and expected shortage of Depot:

$$EBO(s_{i,j}) = \sum_{x=s_{i,j}+1}^{\infty} (x - s_{i,j})P(x|N_{i,j}). \quad (15)$$

**3.2. Mathematical Model of Service Logistics Cost.** The service logistics cost includes the spare parts stocking holding cost as well as maintenance cost and resource allocation cost. In general, maintenance cost and resource allocation cost are often defined in the LORA decision optimization model. If multiple failure modes are considered in service decisions, we find that the steps of iterating LORA model input variables can be simplified.

Maintenance service decisions usually include two kinds of decisions: on-site repair or move to repair. The decision cost can be expressed as  $vc_{i,j,d}$ ,  $d \in \{\text{repair}, \text{move}\}$ . The resource allocation cost can be expressed as  $rc_{r,j}$  ( $j = 1, 2, 3, \dots$ ).

In the case of the single failure model, maintenance service decisions and resource allocation decisions are selected by the method of repeatedly iterating the possibility of all maintenance sites. In the case of multiple failure modes, we believe that different levels of maintenance site service different failure modes and configure different resources. Therefore, the LORA decision model can be simplified as

$$\min \left[ \sum \sum \lambda_{ih,j} \cdot vc_{i,j,\text{repair}} + \sum \sum \lambda_{is,j} \cdot (vc_{i,j,\text{move}} + vc_{i,0,\text{repair}}) + \sum \sum rc_{r,j} \right]. \quad (16)$$

**3.3. Mathematical Model of Spare Parts Stocking Decision Strategy and Service Logistics Cost.** Combining (1) and (16),

we can obtain the spare parts stocking decision strategy and service logistics cost optimization model:

$$\text{s.t.} \begin{cases} \min & \left[ \sum \sum fc_{i,j} \cdot s_{i,j} + \sum \sum \lambda_{ih,j} * vc_{i,j,\text{repair}} + \sum \sum \lambda_{is,j} * (vc_{i,j,\text{move}} + vc_{i,0,\text{repair}}) + \sum \sum rc_{r,j} \right], \\ \text{s.t.} & A(s_{i,j}) \geq A_{\min} \quad (i = 1, 2, 3, \dots; j = 0, 1, 2, 3, \dots), \\ & \lambda_{ih} + \lambda_{is} = \lambda_i \quad (i = 1, 2, 3, \dots). \end{cases} \quad (17)$$

Literature [12] designed an algorithm of splitting and enumerating one by one when solving the two-echelon spare parts stocking decision model. The authors represent the

relationship between components and repair resources with a special graph, in which the vertex represents a resource and the edge represents components. The depth first search

algorithm is used to find the optimal solution. The advantage of this algorithm is intuitive and easy to understand, but the disadvantage is that it needs to be enumerated one by one, and the amount of calculation is huge. Therefore, it is necessary to find a new algorithm to solve the model in this paper.

When solving the lowest spare parts stocking holding cost, a greedy heuristic algorithm is often used [28]. The authors of [10] proposed an approximate method for solving joint optimization, iterative algorithm, but they did not mention spare parts stocking decision optimization, which makes iterative algorithm difficult to solve the model in this paper. Therefore, we improve an iterative greedy heuristic algorithm [23] to solve the optimization model, as shown in Figure 3.

Greedy heuristic algorithm is often used in the decision-making of repairable system spare parts stocking. This algorithm can realize a fast search but can only find a local optimal solution. Iterative method is often used in joint optimization of maintenance level and spare parts allocation strategy. This algorithm traverses the whole world by iterating one by one, but it takes a long time. Combining the advantages of the above two methods, an iterative greedy heuristic algorithm is developed in this study. By iterating spare parts stocking, the local optimal solution of all spare parts allocation strategies is calculated to find the global optimal solution.

The algorithm focuses on balancing spare parts allocation and cost. According to the VAR-METRIC theory [26], we use the marginal analysis method to set the initial stocking of spare parts as 1, increase one stocking step by step until the target availability is met, and then end the algorithm. When the algorithm obtains several groups of Base and Depot stocking combinations, we select the optimal combination to guarantee the system cost.

Step 1: set  $s_{i,j} = 1 + k$ ;  $s_{i,0} = 1 + k$ .

$k$  indicates the number of iterations.  $k = 0$  indicates that the initial stocking level of Base and Depot is 1.

Step 2: set  $\delta_{i,j} = \text{EBO}(s_{i,j}) - \text{EBO}(s_{i,j} + 1)/fc_{i,j}$ ;  $\delta_{i,0} = \text{EBO}(s_{i,0}) - \text{EBO}(s_{i,0} + 1)/fc_{i,j}$ .

$\delta$  is the marginal ability gene. It considers the marginal benefit of adding one item to the stocking [15].

Step 3: select the best value for  $\delta$ . Increase one unit in the stocking of item  $i$  at site  $j$ .

Step 4: calculate the actual availability  $A$ :

if  $A < A_{\min}$ ,  $k = k + 1$ , go to step 2;

if  $A \geq A_{\min}$ , stop the calculation.

Step 5: adjust the stocking parameter and input it into the service logistics cost optimal model.

The iterative process will produce multiple sets of spare parts stocking allocation strategy combinations, and the calculation will stop when the calculated availability is greater than the target availability, which is the local optimal solution of a combination. In this study, the local optimal solutions of all combinations are extracted, and the cost is

calculated. The global optimal solution can be found from the local optimal solution.

Generally speaking, the iterative algorithm can be stopped when the availability satisfies the target availability for the first time. At this moment, the service logistics cost is the optimal cost. In order to be able to traverse the global optimal solution, iteration stops when stocking increases to the number of LRU assemblies ( $M_{i,j}$ ).

## 4. Numerical Experiment

*4.1. Optimization Calculation.* In this section, the energy system of a large-scale laser device is taken as an example to verify the validity of the model and algorithm. The energy system includes 96 homogeneous energy modules. Each module includes three key LRUs: Control Communication module ( $\text{LRU}_1$ ), Gas Switch module ( $\text{LRU}_2$ ), and Trigger module ( $\text{LRU}_3$ ). Due to the different tasks completed by these three LRUs, they were assigned to three different Base sites and managed by the same Depot site, as shown in Figure 4.

These three LRUs will fail randomly. In the study of large-scale system maintenance strategy, literature [29] considered the impact of failure on the delay time of system tasks when distinguishing failure mode. In this example, when distinguishing between hard failure and soft failure, we also consider the impact of failure on system task effect. We attribute the failure that leads to system task cancel to soft failure and the failure that continues to perform tasks through repair at the Base site to hard failure. The specific failure data are shown in Table 2.

Compared with soft failure, hard failure has less impact on system task results and maintenance difficulty. Therefore, the maintenance resource cost of hard failure is lower than that of soft failure. The advantage of soft failure centralized maintenance is to reduce the difficulty and cost of operating environment configuration. To ensure the execution of tasks, large-scale systems have high requirements on the availability of spare parts. We set the minimum availability requirement at 0.98. The specific parameters and cost settings of the optimization model are shown in Table 3.

Spare parts stocking decision cost, maintenance cost, and resource allocation cost of Depot are not comparable with Base. We assume that the transportation costs between the two echelons are the same, and the maintenance resources and components have a one-to-one maintenance relationship. The specific maintenance parameters are shown in Table 4.

According to the iterative greedy heuristic algorithm, we set the initial stocking allocation decision: Base (1, 1, 1) Depot (1, 1, 1). The calculated availability does not meet the requirements obviously. In the iterative process, we can get multiple sets of spare parts stocking allocation combinations. Under the constraints of limited resources, the guaranteed availability calculation results of each combination are shown in Table 5.



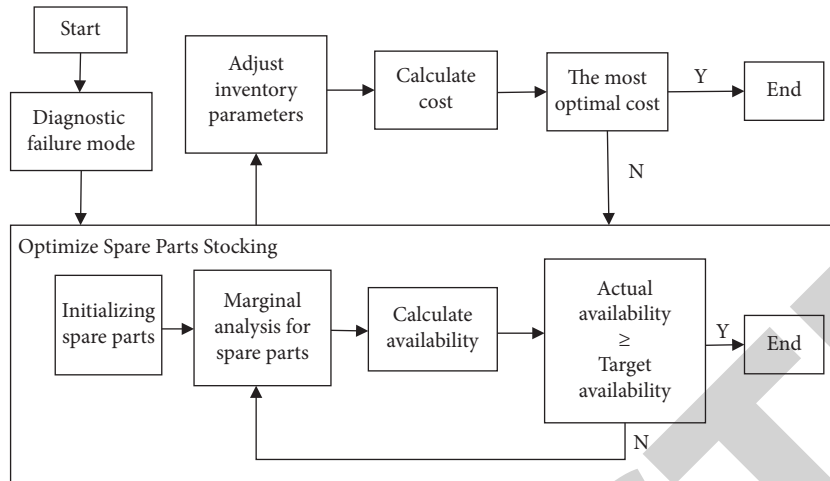


FIGURE 3: The improved iterative greedy heuristic algorithm.

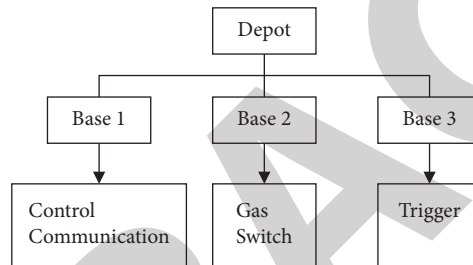


FIGURE 4: Two-echelon service logistics structure of three LRUs.

TABLE 2: LRUs' failure data.

Failure mode	Control communication (LRU <sub>1</sub> )			Gas switch (LRU <sub>2</sub> )			Trigger (LRU <sub>3</sub> )		
	Failure cause	Quantity	Influence	Failure cause	Quantity	Influence	Failure cause	Quantity	Influence
Hard failure	Communication blocking	3	Restart	Communication blocking	3	Restart	Communication blocking	1	Restart
	Software failure	3	Restart	Electromagnetic interference	3	Restart	Electromagnetic interference	1	Restart
	Electromagnetic interference	6	Restart	Barometric instability	2	Delay time	Unstable pilot tube pressure	1	Delay time
Soft failure	PLC controller damaged	3	Cancel	Insulation performance degradation	8	Cancel	Pilot tube performance degradation	10	Cancel
Total		17			16			13	

Extract the three groups of stocking decisions in Table 5 and calculate the service system cost by using formula (17), respectively. The results are shown in Table 6.

We get that the optimal cost is  $214.7 * 10^6$  yuan, and the optimal stocking allocation decision is Base (3, 4, 4) Depot (1, 1, 2). The optimization result conforms to the idea of an iterative greedy heuristic algorithm. That is, when the actual availability meets the minimum requirement, the stocking allocation decision is the optimal decision. Due to the simplified optimization model, the spare parts stocking is the only variable parameter, and the optimal stocking allocation decision guarantees the optimal service logistics cost.

Introducing multiple failure modes into the joint spare parts stocking decision and repair level optimization model can simplify the optimization process and make the optimization results more intuitive and easier to understand.

Combining the results in Tables 4 and 5, it can be found that when the Depot stocking level is (1, 1, 1), no matter how many times the Base stocking iterations are, it cannot satisfy the system availability requirement. This is because the stocking level of Depot plays a decisive role in the system availability. We analyze the relationship between Depot stocking level and system availability in the next section.

TABLE 3: LRUs' parameters.

Parameter cycle: 100 days	Depot		
	Base 1 LRU <sub>1</sub>	Base 2 LRU <sub>2</sub>	Base 3 LRU <sub>3</sub>
$\lambda_i$	17	16	13
$\lambda_{ih}$	12	8	3
$\lambda_{is}$	5	8	10
$T_{i,j}$	0.005	0.01	0.015
$T_{i,0}$	0.01	0.02	0.03
$O_{i,j}$	0.07	0.07	0.07
$fc_{i,j}$ (Base) * 106 × yuan	0.5	0.8	1
$fc_{i,j}$ (Depot) * 106 × yuan	3	5	8
$M_{i,j}$	96	96	192
$A_{\min}$		0.98	

TABLE 4: Maintenance decisions' parameters.

Maintenance decisions	Base 1	Base 2	Base 3	Depot		
	LRU <sub>1</sub>	LRU <sub>2</sub>	LRU <sub>3</sub>	LRU <sub>1</sub>	LRU <sub>2</sub>	LRU <sub>3</sub>
Repair * 10 <sup>6</sup> yuan	1	2	3	3	5	6
Move * 10 <sup>6</sup> yuan	1	1	1	0	0	0
$rc_{r,j}$ * 10 <sup>6</sup> yuan	3	4	5	4	6	8

4.2. *The Relationship between Depot Stocking Level between System Availability.* We select the data in Table 4 and set the horizontal axis as the total stocking level  $S$  of Depot and Base. The vertical axis indicates the system availability, as shown in Figure 5.

Obviously, the improvement of stocking level is very important to improve system availability. In particular, increasing the stocking level of LRU<sub>3</sub> improves the availability significantly.

Because LRU<sub>3</sub> has the largest number of soft failures, it is the most difficult to repair. The optimization model will consider LRU<sub>3</sub> as a key factor affecting decision-making. We can also analyze the importance of the three LRUs under different availability requirements.

4.3. *Importance Analysis of LRUs.* Under different availability requirements, the optimal stocking allocation decisions are shown in Figure 6.

It can be seen that, in the process of improving availability requirements, the ratio of LRU<sub>3</sub> is always the highest. And the ratio of LRU<sub>1</sub> is always the lowest. Compared with LRU<sub>3</sub>, LRU<sub>1</sub> is the factor with the lowest EBO in Depot. Therefore, LRU<sub>1</sub> has the lowest ratio in stocking allocation.

In the lower availability requirement stage (0.80~0.93), the stocking ratio of LRU<sub>2</sub> is the same as LRU<sub>3</sub>. In the higher availability requirement stage (0.93~0.99), the impact of LRU<sub>2</sub> on optimization decision is reduced. This means that as the difficulty of improving availability increases, the allocation of resources concentrates on the most difficult parts to repair.

From the above analysis, it can be seen that (1) the stocking level of the Depot plays a decisive role in ensuring the availability of the system; (2) the LRU with the highest failure rate has the highest influence on the stocking

allocation decisions. In order to improve system availability, repairing time  $T_{i,j}$  is also a key factor that cannot be ignored.

4.4. *Repairing Time Impact Analysis.* We reduce Depot repairing time and Base repairing time, respectively. The above optimization results are reversely iterated. The changes in the availability of stocking allocation decisions are shown in Table 7.

It is obvious that reducing Depot repairing time can improve the availability of service logistics system significantly. At this point, the stocking allocation decision is Base (3, 4, 4) Depot (1, 1, 1). According to the calculation, the service logistics cost of this decision is 206.7 \* 106 yuan, saving 8 \* 106 yuan. However, the reduction of Base repairing time has little effect on the optimization results, which only shows that the availability is improved slightly. From this point of view, Depot repairing time is a key factor that could affect both system availability and service logistics cost.

Reducing Depot repairing time can save a lot of service logistics costs. However, if the Depot's maintenance capacity and stocking resources are limited, the improvement of system availability needs to consider the fronting failure mode. That is, soft failure spare parts are allocated at base site.

4.5. *Fronting Failure Mode.* Combined with Table 4 and Figure 5, we find that when the stocking level of the Depot is (1, 1, 1), no matter how to increase the stocking level of Base, the availability of the service logistics system would never meet the minimum availability requirement.

Fronting failure mode is required.

We consider the front "insulation performance degradation" of LRU<sub>2</sub> and the "pilot tube performance degradation" of LRU<sub>3</sub> into Base site. Their impact on the task of the operating system changes from "cancel" to

TABLE 5: The system availability of different stocking allocation decisions.

Base stocking allocation	Depot stocking allocation			
	(1, 1, 1)	(1, 1, 2)	(1, 2, 2)	(1, 2, 3)
1, 1, 1	0.580867	0.614786	0.625329	0.628775
1, 1, 2	0.699793	0.73096	0.743496	0.74664
1, 2, 2	0.804553	0.840385	0.851464	0.855065
2, 2, 2	0.858257	0.896481	0.908299	0.91214
2, 2, 3	0.898131	0.933406	0.94571	0.949258
2, 3, 3	0.924175	0.960472	0.971907	0.975553
3, 3, 3	0.932103	0.968712	<b>0.980246</b>	<b>0.983922</b>
3, 3, 4	0.940104	0.975766	0.987383	0.99097
3, 4, 4	0.944344	<b>0.980166</b>	0.991567	0.995168
3, 4, 5	0.945577	0.981009	0.992419	0.996024
4, 4, 5	0.94639	0.982047	0.99347	0.997056
4, 5, 5	0.946933	0.98261	0.993996	0.997584

TABLE 6: The service logistics cost of different stocking allocation decisions.

Base stocking allocation			Depot stocking allocation			Service logistics cost * 10 <sup>6</sup> yuan
LRU <sub>1</sub>	LRU <sub>2</sub>	LRU <sub>3</sub>	LRU <sub>1</sub>	LRU <sub>2</sub>	LRU <sub>3</sub>	
3	4	4	1	1	2	214.7
3	3	3	1	2	2	217.9
3	3	3	1	2	3	225.9

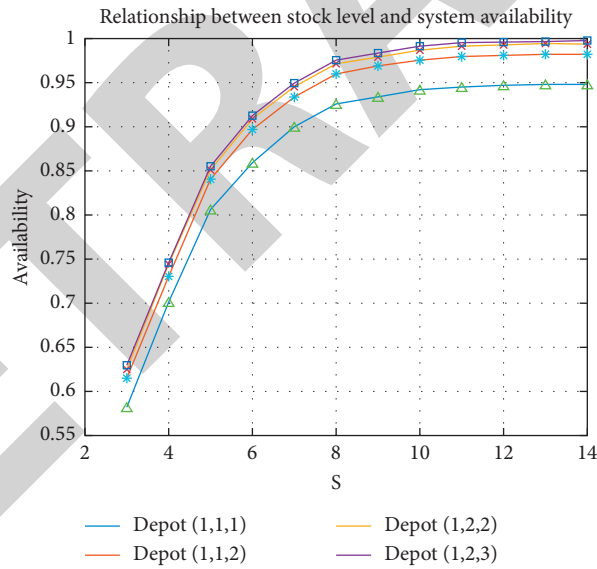


FIGURE 5: Relationship between stocking level and system availability.

“delay time.” The repairing time, spare parts stocking decision cost, and other parameters should be changed accordingly. Fronting failure mode will increase the cost of resource configuration, increase the category of spare parts at the Base maintenance point, and reassign parameters in the optimization model. The specific data are shown in Table 8.

After optimization, the stocking allocation strategy is Base (3, 3, 4) Depot (1, 1, 1), the availability is 0.9803, and the service

logistics cost is 230.9 \* 106 yuan. Thus, if the maintenance capability of the Base could cope with soft failure mode, then the availability of the service logistics system would be improved greatly.

In conclusion, Depot repairing time and Base maintenance capacity are key factors affecting the service logistics system. Reducing Depot repairing time could save service logistics costs; improving Base maintenance capability could improve system availability.

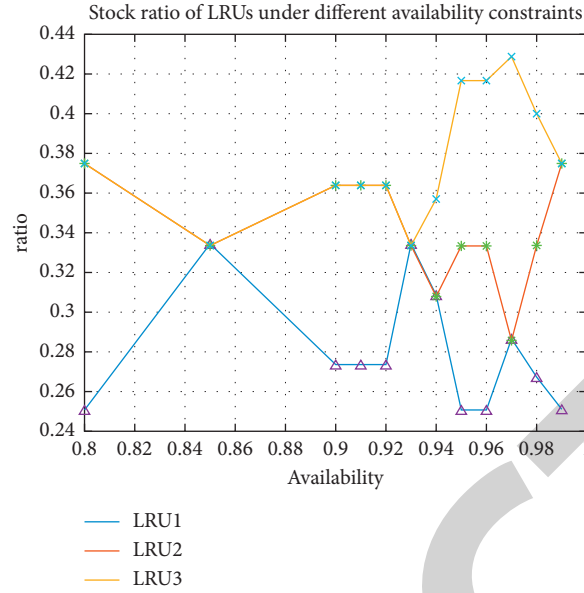


FIGURE 6: Stocking ratio of LRUs under different availability constraints.

TABLE 7: Changes in availability for reducing repairing time.

	Base stocking allocation			Depot stocking allocation			System availability
	LRU <sub>1</sub>	LRU <sub>2</sub>	LRU <sub>3</sub>	LRU <sub>1</sub>	LRU <sub>2</sub>	LRU <sub>3</sub>	
Reduce doubled Depot repairing time	3	4	4	1	1	2	0.993372
	<b>3</b>	<b>4</b>	<b>4</b>	<b>1</b>	<b>1</b>	<b>1</b>	0.983304
	3	4	3	1	1	1	0.976317
Reduce doubled Base repairing time	3	4	4	1	1	2	0.980763
	3	3	4	1	1	2	0.977204
	3	4	3	1	1	2	0.974353

TABLE 8: Parameters after fronting failure mode.

Parameters cycle: 100 days	Base 1	Base 2	Base 3
	LRU <sub>1</sub>	LRU <sub>2</sub>	LRU <sub>3</sub>
$\lambda_i$	17	16	13
$\lambda_{ih}$	12	12	8
$\lambda_{is}$	5	4	5
$T_{i,j}$	0.005	0.015	0.02
$f c_{i,j}$ (Base) * 10 <sup>6</sup> yuan	0.5	1.8	2
Repair * 10 <sup>6</sup> yuan	1	4	5
$r c_{r,j}$	3	6	8

## 5. Conclusion

In this paper, we construct a two-echelon service logistics cost optimization model for large-scale systems. Considering the influence of multiple failure modes on spare parts stocking decisions, the joint optimization model of spare parts stocking decision and repair level was simplified. The iterative greedy heuristic algorithm is improved to find the global optimal solution of stocking allocation strategies. Finally, the validity of the model and algorithm is verified by a practical engineering case.

The contributions and innovations of this paper are as follows: (1) Multiple failure modes are introduced into the joint optimization model to fundamentally simplify the

LORA decisions and reduce the calculation process of the service logistics cost optimization. (2) The improved iterative greedy heuristic algorithm can effectively traverse the global optimal solution. (3) Depot stocking level plays a decisive role in ensuring system availability, and LRU with the highest failure rate has the highest influence on stocking allocation decisions. (4) Depot repairing time and Base maintenance capability are key factors affecting the service logistics system. Reducing Depot repairing time could save costs; improving Base maintenance capability could improve system availability.

As a further improvement, multiple failure modes would lead to discard risk. The maintenance strategy of soft failure could choose imperfect repair. A multiechelon service

logistics system has a potential stocking. The increase of echelons will lead to the increase of stocking iterative combination and the complexity of the algorithm. It can be solved by an improved genetic algorithm [30, 31] and ant colony algorithm [32]. Asset management related to life cycle cost should be considered in maintenance decision-making and spare parts stocking decision management.

## Data Availability

The data used to support the findings of this study are included within the article.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References

- [1] M. S. Mahmoud, "Multilevel systems control and applications: a survey," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 7, no. 3, pp. 125–143, 1977.
- [2] R. Q. Wang, G. Y. Chen, N. Liang, and J. Wu, "Preventive maintenance optimization for deteriorating large-scale systems in life-cycle perspective," *Control and Decision*, 2021.
- [3] Y. H. Yang and Y. C. Feng, "Model research on optimizing allocation of service resources," *Journal of System Simulation*, vol. 13, no. 4, pp. 420–422, 2001.
- [4] R. Basten, M. Heijden, and J. Schutten, "A minimum cost flow model for level of repair analysis," *International Journal of Production Economics*, vol. 133, no. 1, pp. 233–242, 2017.
- [5] G. D. Gravio, F. Costantino, and M. Tronci, "Multi-echelon, multi-indenture spare parts stocking control subject to system availability and budget constraints," *Reliability Engineering & System Safety*, vol. 119, no. 119, pp. 95–101, 2013.
- [6] J. A. Muckstadt, "A model for a multi-item, multi-echelon, multi-indenture inventory system," *Management Science*, vol. 20, no. 4-part-i, pp. 472–481, 1973.
- [7] F. M. Slay, "VARI-METRIC: An Approach to Modeling Multi-Echelon Resupply when the Demand Process Is Poisson with a Gamma Prior," Report AF301-3, Logistics Management Institute, Washington, DC, USA, 1984.
- [8] S. C. Graves, "A multi-echelon stocking model for a repairable item with one-for-one replenishment," *Management Science*, vol. 31, no. 19, pp. 1247–1256, 1985.
- [9] J. W. Rustenburg, G.-J. van Houtum, and W. H. M. Zijm, "Exact and approximate analysis of multi-echelon, multi-indenture spare parts systems with commonality," *International Series in Operations Research & Management Science*, vol. 63, pp. 143–176, 2003.
- [10] R. J. I. Basten, M. C. van der Heijden, and J. M. J. Schutten, "An approximate approach for the joint problem of level of repair analysis and Spare parts stocking decision," *Annals of Operations Research*, vol. 224, no. 1, pp. 121–145, 2015.
- [11] C. H. Falkner, "Jointly optimal stocking and maintenance policies for stochastically failing equipment," *Operations Research*, vol. 16, no. 3, pp. 587–601, 1968.
- [12] R. Basten, M. Heijden, and J. Schutten, "Joint optimization of level of repair analysis and Spare parts stocking decisions," *European Journal of Operational Research*, vol. 222, no. 3, pp. 474–483, 2012.
- [13] Y. H. Wang, C. Deng, X. H. Hu, J. Gao, and C. J. Huang, "Preventive maintenance scheduling for complex mechanical device based on multiple failure mode," *Computer Integrated Manufacturing Systems*, vol. 20, no. 9, pp. 2504–2514, 2015.
- [14] C. Spreafico, D. Russo, and C. Rizzi, "A state-of-the-art review of FMEA/FMECA including patents," *Computer Science Review*, vol. 25, pp. 19–28, 2017.
- [15] Z. P. Ye, Y. Q. Li, X. L. Ma, P. Zhang, and J. X. Lu, "Reliability analysis of EBSM system based on FMEA," *Electronic product reliability and environmental testing*, vol. 38, no. S02, pp. 83–87, 2020.
- [16] L. He, Z. Lu, and X. Li, "Failure-mode importance measures in structural system with multiple failure modes and its estimation using copula," *Reliability Engineering & System Safety*, vol. 174, pp. 53–59, 2018.
- [17] R. Mullor, J. Mulero, and M. Trovati, "A modelling approach to optimal imperfect maintenance of repairable equipment with multiple failure modes," *Computers & Industrial Engineering*, vol. 128, pp. 24–31, 2019.
- [18] H. Dui, X. Meng, H. Xiao, and J. Guo, "Analysis of the cascading failure for scale-free networks based on a multi-strategy evolutionary game," *Reliability Engineering & System Safety*, vol. 199, Article ID 106919, 2020.
- [19] *Failure Modes and Effects Analysis FMEA*, Springer International Publishing, Detroit, MI, USA, 2004.
- [20] R. Peng, X. F. He, C. Zhong, G. Kou, and H. Xiao, "Preventive maintenance for heterogeneous parallel systems with two failure modes," *Reliability Engineering & System Safety*, vol. 220, Article ID 108310, 2022.
- [21] C. Duan, C. Deng, Q. Gong, and Y. Wang, "Optimal failure mode-based preventive maintenance scheduling for a complex mechanical device," *International Journal of Advanced Manufacturing Technology*, vol. 95, pp. 2717–2728, 2018.
- [22] S. Taghipour and D. Banjevic, "Optimal inspection interval for a system under periodic and opportunistic inspections," *IIE Transactions*, vol. 44, no. 11, pp. 932–948, 2012.
- [23] S. Taghipour and D. Banjevic, "Optimal inspection of a complex system subject to periodic and opportunistic inspections and preventive replacements," *European Journal of Operational Research*, vol. 220, no. 3, pp. 649–660, 2012.
- [24] W. Zhu, M. Fouladirad, and C. Brenguer, "A multi-level maintenance policy for a multi-component and multi-failure mode system with two independent failure modes," *Reliability Engineering & System Safety*, vol. 153, pp. 50–63, 2016.
- [25] Q. Huang, "Simulation model of multi-echelon closed supply chain for spare parts," in *Proceedings of the IEEE 2011 3rd International Workshop on Intelligent Systems & Applications (ISA)*, pp. 1–4, Wuhan, China, May 2011.
- [26] C. C. Sherbrooke, "METRIC: a multi-echelon technique for recoverable item control," *Operations Research*, vol. 16, no. 1, pp. 122–141, 1968.
- [27] C. C. Sherbrooke, *Optimal Stocking Modeling of Systems*, pp. 101–125, Multi-echelon techniques, Kluwer, Dordrecht, 2nd edition, 2005.
- [28] H. Zhang, X. L. Hua, and S. J. Xu, "Simulation optimization model of stocking decision for repairable spares systems," *Systems Engineering and Electronics*, vol. 31, no. 6, pp. 1510–1514, 2009.
- [29] D. Lin, M. Zuo, and R. Yam, "Sequential imperfect preventive maintenance models with two categories of failure modes," *Naval Research Logistics*, vol. 48, no. 2, pp. 172–183, 2001.
- [30] H. Sadeghpour, A. Tavakoli, M. Kazemi, and A. Pooya, "A novel approximate dynamic programming approach for constrained equipment replacement problems: a case study," *Advances in Production Engineering & Management*, vol. 14, no. 3, pp. 355–366, 2019.

## Retraction

# Retracted: A Prediction Model of Health Development Based on Linear Sequential Extreme Learning Machine Algorithm Matrix

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] S. Cheng and S. Liu, "A Prediction Model of Health Development Based on Linear Sequential Extreme Learning Machine Algorithm Matrix," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 7632841, 12 pages, 2022.



## Research Article

# A Prediction Model of Health Development Based on Linear Sequential Extreme Learning Machine Algorithm Matrix

Suli Cheng<sup>1</sup> and Shuzhi Liu <sup>2</sup>

<sup>1</sup>College of Mathematics and Statistica, Chongqing Technology and Business University, Chongqing 400067, China

<sup>2</sup>School of Physical Education, Chongqing University, Chongqing 400044, China

Correspondence should be addressed to Shuzhi Liu; [cqutyliushuzhi@cqu.edu.cn](mailto:cqutyliushuzhi@cqu.edu.cn)

Received 24 December 2021; Revised 12 February 2022; Accepted 15 February 2022; Published 7 March 2022

Academic Editor: Daqing Gong

Copyright © 2022 Suli Cheng and Shuzhi Liu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The rapid development of social economy not only increases people's living pressure but also reduces people's health. Looking for a healthy development prediction model has become a domestic concern. Based on the analysis of the influencing factors of health development, this paper looks for a model to predict the development of public health, so as to improve the accuracy of health development prediction. In this paper, the linear sequential extreme learning machine algorithm can be used to evaluate the health status of a large number of data, analyze the differences of each evaluation index, and construct the analysis model of health status. Therefore, this paper introduces rough set theory into linear sequential extreme learning machine algorithm. Rough set can analyze the double analysis of evaluation scheme, predict the health development of different individuals, and improve the evaluation accuracy of mass health evaluation. The simulation results show that the improved line sequential extreme learning machine algorithm can accurately analyze the mass health and meet the needs of different individuals' health evaluation.

## 1. Introduction

With the acceleration of the pace of life and the increase of people's work pressure, people pay more and more attention to their health. People pay attention to their own health. Health is divided into subhealth, health, and various diseases. However, due to the individual differences, basic diseases, and their own health quality, the health judgment scheme cannot effectively meet the needs of the public [1]. The survey results show that by 2020, the number of patients with health abnormalities in China will reach 2232/100000, and by 2030, this data will increase by 30%. Among them, 20% of the underground people are unable to take corresponding measures in time due to unreasonable prediction. Therefore, it is an urgent problem to formulate an effective judgment scheme according to personal constitution. Local governments, society, and nongovernmental organizations regularly monitor their health to let the public know their health [2]. At the same time, local governments should also elaborate on the methods of health prediction and publicize

them through television, Internet, and other channels. At present, the key of health prediction is how to evaluate relevant indicators and guide different individuals to carry out health test [3]. Some scholars believe that the formulation of health prediction scheme should be combined with personal situation, carry out multiangle regression analysis, and determine the impact of different factors on the test results. Some scholars also believe that the reasonable formulation of health prediction scheme and test plan will enable the public to more accurately understand their health condition; otherwise, it will affect the public's health. At present, there are many methods to analyze mass health, mainly including Bayesian method, genetic algorithm, time series algorithm, and differential evolution algorithm based on vague local search strategy [4]. However, the above algorithm only makes a unified analysis of mass health and cannot realize the differential analysis of different bodies. In order to solve this problem, some scholars put forward the health prediction model based on the combination of rough analysis method and linear sequential extreme learning

machine and analyzed the mass health index by using rough set. Rough set method is used to couple the eigenvalues of the data in the health index, and the eigenvalues are extracted by threshold [5], penalty coefficient, weight coefficient, and mean value to obtain more accurate pre-processed data. It can be seen that the linear sequential extreme learning machine algorithm has been applied in the healthy development, but there are few application cases, which need to be further analyzed. Linear sequential extreme learning machine method has the advantage of continuous prediction of healthy development, and rough set has the advantage of massive data processing. The combination of the two can carry out continuous analysis of healthy development prediction. The difference analysis of linear sequential extreme learning machine and its own iterative operation can obtain the optimal set of health data. Among them, the constraint conditions and constraint coefficients can avoid the local extremum problem and improve the efficiency of the calculation results [6]. To sum up, domestic scholars' research on health prediction and analysis mainly focuses on the improvement of indicators, ignoring the coupling between indicators and the rationality of prediction scheme. Therefore, the prediction of healthy development urgently needs a model that can always carry out coupling analysis of various indicators. However, there is less research on continuous prediction of healthy development in China, and there is less research on the application of rough set and linear sequential extreme learning machine. Compared with China, there are many researches on online sequential extreme learning machine abroad, and it is combined with k-clustering and rough set to evaluate the healthy development and achieved good results. So, it is urgent to improve the linear sequential extreme learning machine and apply it to the prediction of healthy development. On this basis, in order to find a more effective model, the improvement of line sequence extreme learning machine and its application to healthy development prediction have become the key to solve the above problems. In addition [7], a single analysis of mass health will not only reduce the accuracy of evaluation results but also affect the future prediction of health. In addition, although the linear sequential extreme learning machine can carry out iterative analysis and predict the mass health, it cannot analyze the individual health. Some scholars believe that the combined analysis of personality and unity can promote the development of health prediction and play a positive role in the improvement of health [8]. Therefore, by optimizing the line sequential extreme learning machine algorithm and combining the rough set processing method, we can accurately evaluate the mass health and obtain the Pareto optimal set. Based on the above theoretical analysis, this paper proposes an improved line sequential extreme learning machine to study mass health [9]. The specific content includes three aspects: the first part introduces the research status of rough set line sequential extreme learning machine and mass health. In the second part, an improved line sequential extreme learning machine algorithm is constructed, and the numerical judgment and constraints of rough set are explained. The third part verifies and analyzes

the improved line sequential extreme learning machine algorithm to judge the accuracy and time of mass health calculation [10]. Compared with other studies at home and abroad, this paper optimizes the line sequential extreme learning machine algorithm and combines this method with rough set to obtain the Pareto optimal set under the constraint coefficient [11]. Through the joint analysis of unified health judgment and individual health, we can realize the dual verification of the above two aspects, improve the accuracy of the analysis results, promote the reasonable construction of health evaluation scheme, and finally achieve the purpose of improving mass health.

## 2. The Algorithm Description Based Online Sequential Extreme Learning Machine

The linear sequential extreme learning machine algorithm was first applied to the diagnosis of space engine. It can predict the abnormal situation of aircraft through equal weight processing of data. However, the algorithm cannot realize the analysis of massive data, and the speed of processing massive data is slow. Rough set is a set proposed by Pawlak Z. It can comprehensively process incomplete data through induction, learning, and mining to build a relatively clear and concise data system to support subsequent composition analysis. Based on the above analysis, this paper uses rough set theory to collect mass physical health data [12], supplement incomplete data and indicators, and finally get a clearer data system. The specific data processing flow is shown in Figure 1.

As can be seen from Figure 1, the first step is to obtain physical health data, analyze the data types, and classify them according to the types. The second step is to formulate physical health plan or adjust the plan according to the analysis results of physical health data. The third step is to test the physical health of the masses, compare the effects of different exercise methods, and record the matching results between the exercise scheme and the masses. The fourth step is to record the optimal matching scheme into the database and eliminate the mismatched evaluation scheme. Medical institutions, rehabilitation centers, elderly care centers, sports centers, and fitness centers in four provinces were observed, and preliminary unified data were obtained by means of video recording, heartbeat detection, and vital capacity test. Combined with relevant domestic literature [13], 23 physical health indexes and 12 physical health evaluation indexes are obtained. After experts' determination and 120 sampling questionnaires, there is no significant correlation between the above indexes, which can be used as evaluation indexes and samples. The visit sites are located in the northeast, northwest, southeast, and southwest [14], as well as north and central China. The access data come from online questionnaires, actual data collection, official published materials, and other internal materials. The time span of accessing data is 2–3 years, and the personnel span is more than 4. The validity and reliability of the survey results were >0.7. Finally, specific evaluation indicators are obtained, as shown in Table 1.

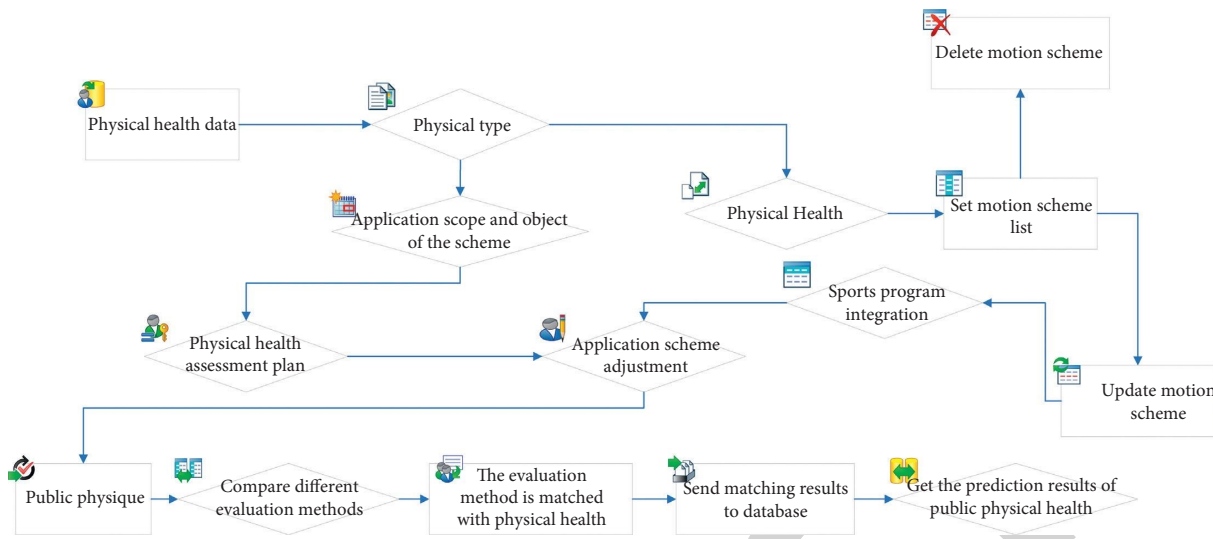


FIGURE 1: The data processing process of physical health development prediction.

TABLE 1: The mass physique and health evaluation index.

The index category		The quantity	The correlation with other indicators
Amount of exercise	Type of motion	4	0.232
	Exercise time	3	0.767
	Exercise intensity	2	0.562
Body function	Blood oxygen content	9	0.222
	Respiratory rate	4	0.721
	Cardiac blood supply	6	0.492
	Lung volume	3	0.662
	Inflammatory factor index	2	0.892

Note. There is no significant correlation between the indicators, and the reliability and validity are >0.7. The data come from online literature and actual survey.

### 3. The Prediction Model of Physical Health Development Based on Linear Sequential Extreme Learning Machine Algorithm Matrix Is Established

3.1. Improved Line Sequence Extreme Learning Machine Algorithm Matrix. First, add rough sets. In the physical health analysis, the physical health database, physical prediction database, and their correlation database are established by rough set method. Then, rough set uses the above database to analyze the algorithm matrix of linear sequential extreme learning machine and determine the mass constitution. In the process of analysis [15], the weight of different physical health prediction schemes is determined through the correlation, dependence, and correlation between physical health prediction schemes and mass physique. Due to individual differences, basic diseases, knowledge level, and other factors, there are deviations in the impact of the prediction scheme on the prediction results of physical health. Rough set regulates the calculation direction of feature data by setting penalty value, avoids the local extremum problem of line sequential extreme learning machine algorithm matrix, and improves the accuracy of calculation results.

Second, add discrete clustering. The data in rough set present discrete state. Although the weight coefficient is set, it is still unable to realize the continuous analysis of data. How to use finite eigenvalues and feature points to construct continuous rough set sequences is the key of this paper [16]. In this paper, rough set is introduced to classify eigenvalues, and discrete data sequences are constructed by using the correlation and complementarity between adjacent features. At the same time, rough set uses the concept of fuzzy mathematics to eliminate the eigenvalues with high dispersion, calculate the membership relationship between eigenvalues more accurately, and realize the data optimization of linear sequential extreme learning machine algorithm matrix. The improvement process of linear sequential extreme learning machine algorithm matrix is shown in Figure 2.

According to the analysis in Figure 2, rough set analysis is carried out on structured and semistructured data to realize data coupling, fusion, and eigenvalue extraction. Rough set realizes the preprocessing of mass health data. Rough set realizes the orderly arrangement of characteristic data, improves the accuracy of data processing, reduces the data processing time, and increases the amount of pre-processed data.

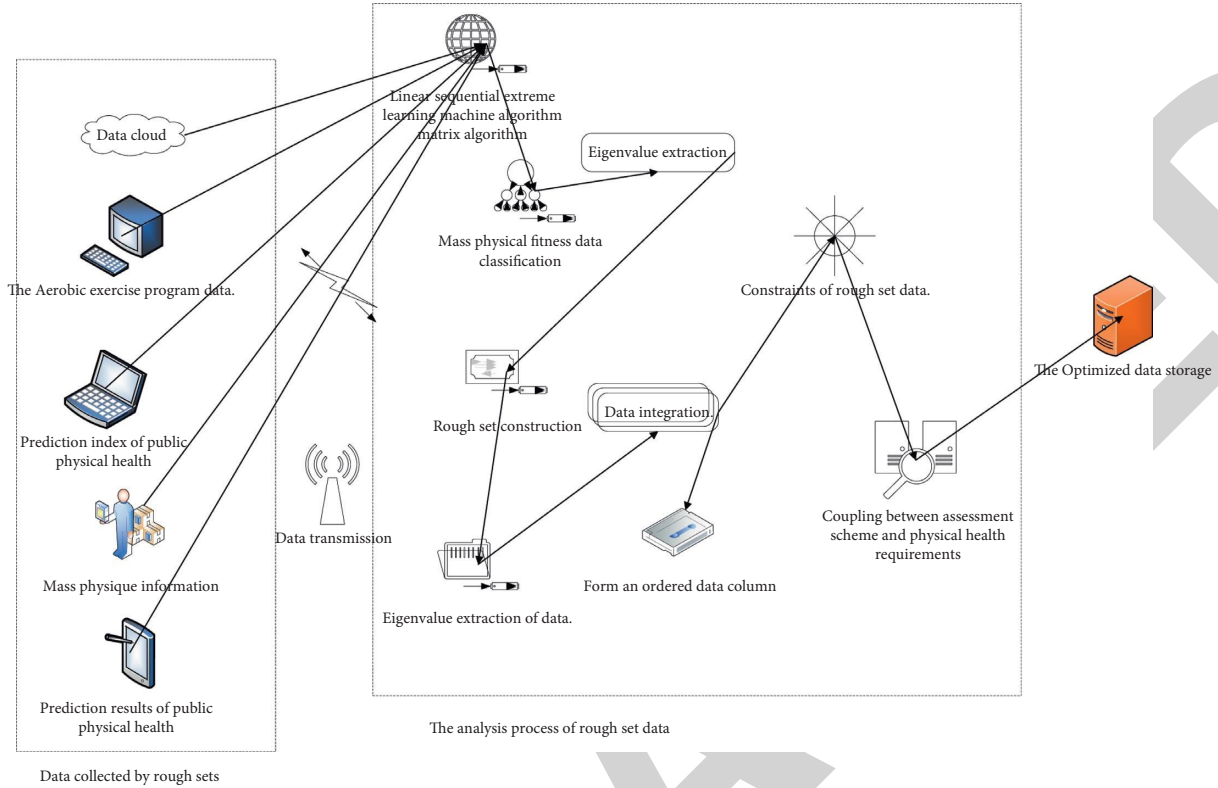


FIGURE 2: The physical health assessment scheme and prediction.

**3.2. Analysis of Mass Physical Health Evaluation Scheme Based on Linear Sequential Extreme Learning Machine Algorithm Matrix.** In the case of individual differences, the physical health evaluation scheme has a significant impact on the evaluation results, and there are significant differences in the effects of blood oxygen index, blood pressure, and heart rate [17]. Although the linear sequential extreme learning machine algorithm matrix can carry out progressive analysis on mass physical health data, the analysis results are discrete. Rough set can cluster the discrete results, submit the accuracy of the results, and arrange the eigenvalues of the results in order. Rough set can double analyze the physical health scheme and physical data, so as to further improve the accuracy of the results. Therefore, the combination of linear sequential extreme learning machine algorithm matrix and rough set can orderly arrange the characteristic data, improve the accuracy of data analysis results, and realize the double evaluation of mass physical health. The specific evaluation scheme is shown in Figure 3.

It can be seen from Figure 3 that different masses adopt different physical fitness assessment scheme, and different schemes will also have different effects on physical health. Therefore, we should adjust the physical fitness assessment

scheme according to the individual differences of masses, so as to achieve the best regulation of physical health.

### 3.3. The Algorithm Construction

**3.3.1. Objective Function Construction of Physical Health Evaluation.** Assuming that the evaluation objective of physical health is  $X$ ,  $y_i$  is the physical fitness assessment scheme,  $i$  is the exercise program number, and  $j$  is the correlation between the physical fitness assessment scheme and physical health,  $X_{ij}$  is the evaluation objective of any physical health; the calculation results are shown in

$$\min X = \sum_{i,j}^T f[x_{ij}(1 - x_{ij-1})]S_{ij} + \zeta(y_{ij}), \quad (1)$$

where  $S_{ij}$  is the implementation effect of  $j$  exercise program,  $\zeta(y_{ij})$  is the impact rate of  $j$  exercise program on physical health, and  $T$  is the implementation process of physical fitness assessment scheme.  $\zeta(y_{ij})$  is the influence rate of the evaluation scheme; the calculation results are shown in

$$\zeta(y_{ij}) = \alpha_i + \beta_i y_{ij-1} + \chi_i y_{j-1}^2, \quad (2)$$

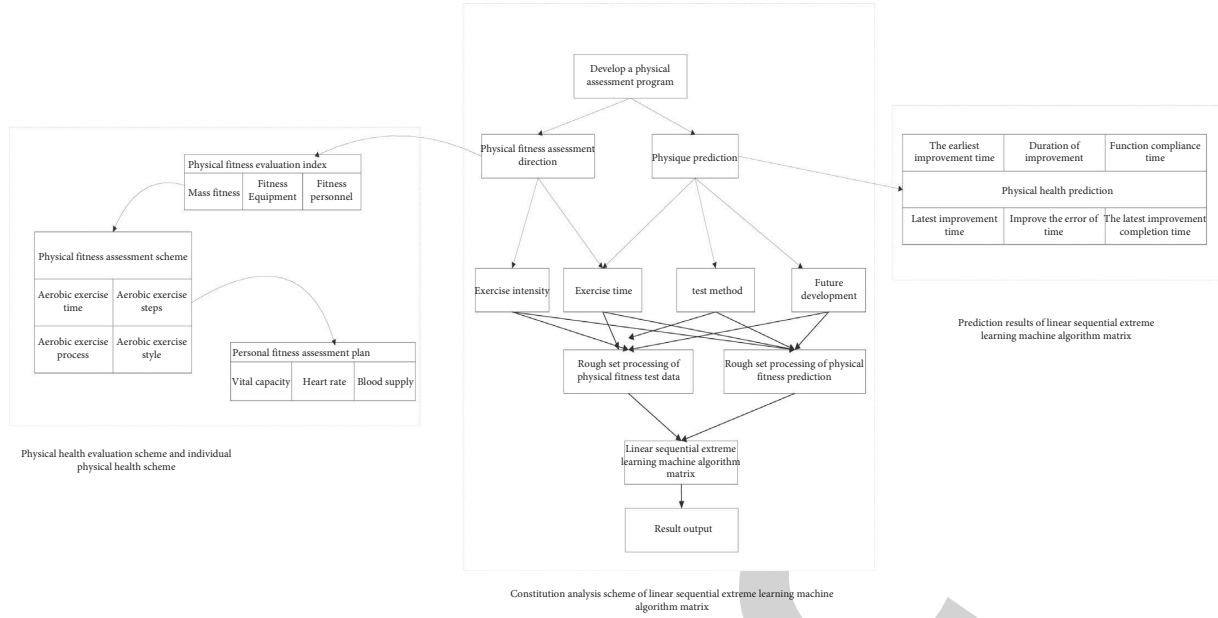


FIGURE 3: The implementation process of linear sequential extreme learning machine algorithm matrix.

where  $\alpha_i$ ,  $\beta_i$ , and  $\chi_i$  are the influence coefficients of the evaluation scheme.

### 3.3.2. Evaluation Function Construction of Aerobic Regimen.

It is assumed that the evaluation  $H$  of aerobic scheme is in three states: scientific and reasonable state  $a_i$ , vague state  $b_i$ , and negative influence state  $d_i$ ; the calculation results are shown in

$$H(y_{ij}) = \prod_{i=1}^T \sum_{i=1}^T A_{ai} + B_{bi} + Cc_i. \quad (3)$$

Among them,  $A$ ,  $B$ , and  $C$  are the evaluation coefficients, and  $T$  is the time node of different evaluation schemes. The evaluation coefficient is mainly determined by the mass health association, aerobic fitness association, and Chinese Medical Association. At the same time, the evaluation coefficient of each region can be adjusted according to the statistical data of 10 years, and the adjustment proportion is less than 0.3.

3.3.3. Construction of Constraint Function. It is assumed that the upper limit of the evaluation result of physical fitness assessment scheme is  $L_{max}$  and the lower limit is  $L_{min}$ . The upper limit of physical health evaluation was  $l_{max}$ , and the lower limit was  $l_{min}$ . The calculation results are shown in

$$Z(X, H) = \sum_{\min=0}^{\infty} \frac{\max(X, H) - (A_{ai} + B_{bi} + Cc_i)'}{(A_{ai} + B_{bi} + Cc_i)' - \min(X, H)} \cdot \prod_{\min} f[x_{ij}(1 - x_{ij-1})]S_{ij}', \quad (4)$$

where  $\max(X, H)$  is the maximum value set of evaluation results,  $\min(X, H)$  is the minimum value set of evaluation

results, and  $(A_{ai} + B_{bi} + Cc_i)'$  is any maximum value.  $S_{ij}'$  is optimal physical fitness assessment scheme. The specific construction process is shown in Figure 4.

It can be seen from Figure 4 that the calculation process of physical fitness assessment scheme and physical health prediction exceeds the expected maximum and minimum values, indicating that the setting of constraints can limit the differential evolution process of vague local search strategy. At the same time, the physical fitness assessment scheme and the improvement of physical health did not change significantly, showing regular ups and downs. At the same time, the following picture in Figure 4 shows a large depression, indicating that the data here are an inflection point. However, the data in the depression still changes smoothly, which further shows that the data processing effect is better.

3.4. The Construction of Linear Sequential Extreme Learning Machine Algorithm Matrix. The linear sequential extreme learning machine algorithm matrix based on vague local search strategy is mainly in three aspects. The calculation results are shown in

$$|E| = Z(\cdot) \times \begin{cases} \sum_{i,j=0}^n \frac{x_{ij}, y_{ij}}{\dots} \\ \sum_{i=0}^n S_j^{H((2\pi/2)(x_{ij}, y_{ij}))} \end{cases}, \quad (5)$$

where  $x_{ij}$  and  $y_{ij}$  are the approximate set of  $x_{ij}$  and  $y_{ij}$ , and  $|E|$  is the coupling set.  $i \neq j$ ,  $i, j \in \{1, \dots, n\}$ ,  $\cup_{i,j=1}^n X_{ij}, H_{ij} = U$ .  $U$  is the correlation between  $X_{ij}$  and  $H_{ij}$ . The second aspect is to cluster discrete eigenvalues by using rough set method to reduce the discreteness of eigenvalues. The specific formula is as follows.



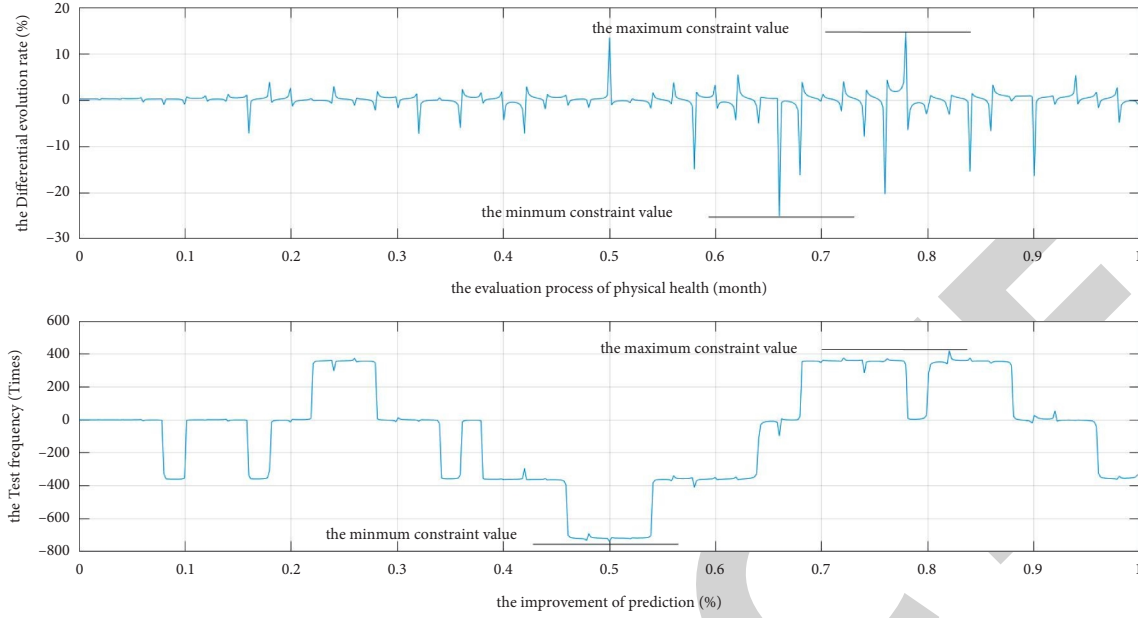


FIGURE 4: The constraint analysis of physical fitness assessment scheme and physical health prediction.

Assuming the sample set  $X = \{x_{11}, x_{12} \cdots x_{ij}\}$ ,  $H = \{h_{11}, h_{12} \cdots h_{ij}\}$ , the clustering center  $V$  of each dimension is shown in

$$V = \{v_{11}, v_{12} \cdots v_{ij}\}. \quad (6)$$

The membership matrix  $u$  of each cluster center  $V$  is shown in the formula. The calculation results are shown in

$$u = \begin{bmatrix} u_{11} & \cdots & \cdots & u_{1j} \\ \vdots & \ddots & & \vdots \\ u_{i1} & & & u_{ij} \end{bmatrix}, \quad (7)$$

where  $u_{ij}$  is the subordinate relationship between the sample  $u_i$  and  $u_j$ . The formula of rough set function is shown in

$$\min F(u_{ij}, v_i) = \sum_{i=1}^{\tau} \sum_{j=1}^{\tau} u_{ij}^{\tau} (x_{ij} - v_i), \quad (8)$$

where  $\tau$  is the clustering index between  $X_{ij}$  and  $H_{ij}$ . The calculation of  $v_i$  cluster center is shown in

$$v_i = \frac{x_{ij}}{\sum_{i=1}^{\tau} u_{ij}^{\tau} (x_{ij} - v_i)} \sum u_{ij}^{\tau} (x_{ij} - v_i). \quad (9)$$

The calculation of membership relationship is shown in

$$u_{ij} = \frac{(1/\Delta x_{ij})^{1/(\tau-1)}}{\sum_{i=1}^{\tau} [1/(\Delta x_{ij} - v_i)]^{1/(\tau-1)}}. \quad (10)$$

The third aspect is to determine the weight coefficient. In order to more accurately analyze the physical fitness assessment scheme and realize the comprehensive evaluation of the impact of physical health [18], the weight of each index should be assigned. At the same time, the weight assignment

can avoid the local extreme value problem and improve the accuracy of the evaluation results. Firstly, each objective function is regarded as a conditional attribute [19], its weight is obtained, and the optimal template function  $X$  value is calculated. Secondly, the evaluation template attribute set is obtained [20]. The optimal scheme obtained under different physical fitness assessment scheme is  $h$ ; then, any physical fitness assessment scheme in  $X$  is  $u_{ij}$ .

Therefore, the dependence between physical fitness assessment scheme and physical health evaluation is shown in the formula. Firstly, each objective function is regarded as a conditional attribute,  $w_i = 1/n_i$  weight is obtained, and the optimal template function  $X$  value is calculated. Secondly, the evaluation template attribute set  $D = \{X\}$  is obtained. The optimal scheme obtained under different physical fitness assessment scheme is  $H$ ; then, any physical fitness assessment scheme in  $X$  is  $u_i$ .

Therefore, the dependence between physical fitness assessment scheme and physical health evaluation is shown in

$$r_i(H) = \frac{\sum \rho[r_i[F]_D]}{\rho(u_i)}. \quad (11)$$

The degree of dependence between evaluation objectives is shown in

$$\Delta r_i(H) = \frac{\sum \rho[\Delta r_i[F]_D]}{\rho(\Delta u_i)}. \quad (12)$$

The importance of the  $i$ th physical health assessment is shown in

$$\sigma(X, H) = r_i(H) - \Delta r_i(H). \quad (13)$$

The weight between the  $i$ th physical health assessment and physical health is shown in



$$w_i = \sigma(X_i, H_i) / \sum \sigma(X, H). \quad (14)$$

In order to give weight more accurately, increase the residual value of weight change. The calculation results are shown in

$$w_i = \sigma(X_i, H_i) / \sum \sigma(X, H) + \xi, \quad (15)$$

where  $\xi$  is the residual value of weight, which is mainly determined by the development of physical health in different regions.

**3.5. The Output of Physical Health Evaluation Results under Linear Sequential Extreme Learning Machine Algorithm Matrix.** Firstly, the physical fitness assessment scheme output of different individuals: physical health is multidimensional, complex, and irregular, so it is affected by individual physical quality, basic diseases, training methods, training time, and other factors.

Assuming that different influence factors are  $k_i$ , the output of physical fitness assessment scheme under different factors is shown in

$$H(y_{ij})_k = v_i \cdot \prod_k \lambda \cdot \sum_{i=1}^T A_{ai} + B_{bi} + Cc_i, \quad (16)$$

$v_i$  is the data after rough set,  $k$  is the influencing factor,  $H(y_{ij})_k$  is the output optimal physical fitness assessment scheme, and  $\lambda$  is the differential evolution coefficient. Therefore, the application of rough set method improves the adaptability of physical fitness assessment scheme, meets the exercise needs of different individuals, and realizes the effective aggregation of discrete eigenvalues.

**3.6. Physical Health Evaluation Output of Different Physical Fitness Assessment Schemes.** Different physical fitness assessment schemes have different effects on physical health, so it is necessary to accurately evaluate physical health. Generally speaking, the improvement of physical health is judged in the form of ten-point system. Under the action of rough set, the output results of physical health evaluation are calculated. The calculation results are shown in

$$X_{H_i} = \frac{w_i \cdot \varsigma(y_{ij}) \sum_{i,j}^T f[x_{ij}(1 - x_{ij-1})] S_{ij} + \varsigma(y_{ij})}{v_i \cdot \prod_k \lambda \cdot \sum_{i=1}^T A_{ai} + B_{bi} + Cc_i}, \quad (17)$$

where  $H_i$  is different physical fitness assessment scheme and  $\lambda$  is the coefficient of differential evolution.

Secondly, the calculation steps of physical health evaluation results. Through the above analysis, the overall overview of physical health analysis can be obtained. The specific steps are as follows.

First, construct a rough set of the impact of such exercise on physical health, and form  $X = \{x_{11}, x_{12} \cdots x_{ij}\}$  and  $H = \{h_{11}, h_{12} \cdots h_{ij}\}$  sets.

Secondly, the data set is constrained, the clustering center  $V$  of the data is calculated, the weights of physical fitness assessment scheme and physical health are obtained,

and the corresponding iterative operation is carried out. If the calculation results  $X$  and  $H$  are greater than  $\max\{X, H\}$ , the results shall be included in the calculation results; otherwise, they shall be eliminated. If the calculation result is less than  $\min\{X, H\}$ , the result shall be included in the calculation result; otherwise, it shall be eliminated.

Third, when all 35 indicators are traversed, output  $H(y_{ij})_k$ ,  $X_{H_i}$ ; otherwise, repeat Step 2.

## 4. The Cases of Physical Health Assessment and Prediction

**4.1. The Research Data through the 35 Indicators Obtained from the Survey: A Case Study Is Carried Out.** The data in the case mainly comes from 6 randomly selected cities, and the verification method is mainly to compare before and after (note: the validity and reliability of the data are  $>0.7$ ) [21]. At the same time, the linear sequential extreme learning machine algorithm matrix is compared with the original algorithm (note: linear sequential extreme learning machine algorithm is not combined with rough set). The verification period is January, March, June, and December, and the exercise intensity is 60 H/week. The subjects were 25–65 years old, with an average age of  $32.2 \pm 2.32$  years [22]; BMI 23–40, average BMI  $28.2 \pm 1.12$ . Inclusion criteria: (1) Meet the 1994 edition of American diagnostic criteria for physical health. (2) The subjects had no previous physical health and circulatory diseases. (3) Get the consent of the ethics committee of the hospital and sign the informed consent form. (4) The tester can express himself. (5) The tester has some fitness experience. The results are shown in Table 2.

The data in Table 2 are the evaluation contents in different test time periods, mainly including training intensity and training score, and the data are from the experimental results and network literature. In order to ensure the accuracy of the test, the tester should be analyzed before the test. According to different methods, they were divided into improvement group and original group. The improved group adopts the differential evolution algorithm of improved vague local search strategy, and the original group adopts the differential evolution algorithm of vague local search strategy. The results are shown in Table 3.

It can be seen from Table 3 that there is no significant difference between the two groups in physical health, pulmonary respiratory capacity, blood pressure, quality of life, blood oxygen content, and pulse, so it can be compared and analyzed.

**4.2. Evaluation of Physical Fitness Assessment Scheme by Linear Sequential Extreme Learning Machine Algorithm Matrix.** The accuracy of physical fitness evaluation scheme evaluation is the key of evaluation. The linear sequential extreme learning machine algorithm improves the fuzzy local search strategy. The specific evaluation results are shown in Figure 5.

It can be seen from Figure 5 that the average accuracy of physical fitness assessment scheme is more than 90%, and the effective scheme is more than 60%. At the same time, the

TABLE 2: The evaluation contents of test.

Duration	1~2 h	2~2.3 h	1~3.2 h	1~2 h	1~1.5 h	1~1.5 h	1~2 h	1~3 h
Exercise time	300 KJ	4333 KJ	1222 KJ	4333 KJ	2392 KJ	1223 KJ	4303 KJ	2322 KJ
Exercise intensity	1-5	7-11	9-22	11-19	33-89	102-111	47-82	46-90
Physical health score	10 score	10 score	10 score	10 score	10 score	10 score	10 score	10 score
Physical health degree	3	1	3	2	1	1	2	1

TABLE 3: The comparison of pretest indexes of masses [n,  $\bar{x} \pm s$ ].

The group	Physical health score (score)	Physique index (piece)	Blood pressure (cm H <sub>2</sub> O)	Physical fitness standard rate (%)	Quality of life score (m3/ml)	Pulse (mmHg)
Improvement group ( $n = 20$ )	$6.33 \pm 5.21$	$28.61 \pm 6.42$	$78.34 \pm 5.27$	$88.71 \pm 6.22$ *	$186.49 \pm 11.24$ *	$70.13 \pm 9.82$ *
Original group ( $n = 20$ )	$6.84 \pm 5.32$	$28.97 \pm 6.53$	$78.52 \pm 5.42$	$87.42 \pm 6.97$ *	$179.14 \pm 7.27$ *	$73.62 \pm 8.14$ *
$T$	0.907	0.176	0.107	0.192	0.137	0.789
$P$	0.370	0.861	0.916	0.893	0.726	0.822

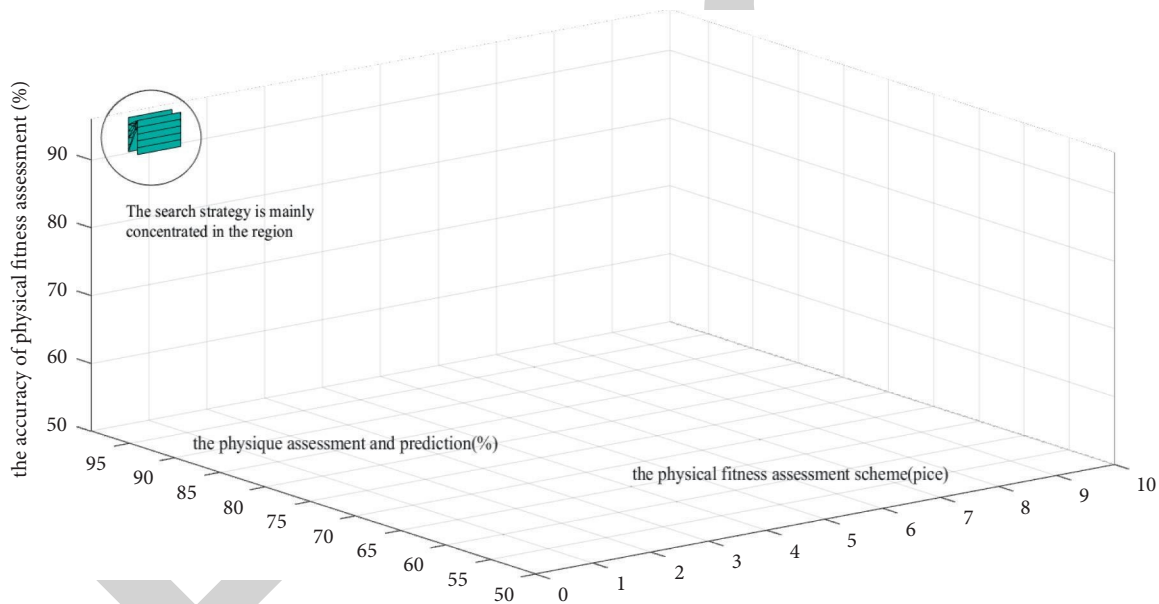


FIGURE 5: The evaluation results of physical health.

effective program of physical health was the highest at 96% and then decreased. The results of this study are consistent with relevant domestic studies. The reason is that after the human physical health reaches the maximum limit, the body's self-protection mechanism starts, resulting in a decline in the improvement of physical health. The linear sequential extreme learning machine algorithm matrix is compared with the original algorithm, and the results are shown in Figure 6.

It can be seen from Figure 6 that the linear sequential extreme learning machine algorithm matrix has high evaluation accuracy of physical health, greater than 99%, while the accuracy of the original algorithm is greater than 90%; the results of this study are consistent with relevant domestic studies. This shows that the linear sequential extreme learning machine algorithm matrix is better in physical fitness assessment scheme. It can be seen from the figure that

the evaluation results of the two methods change smoothly. However, the model proposed in this paper changes straightly, while the original line sequence extreme learning machine algorithm changes at the positions of the 4th, 6th, 11th, 13th, 16th, 22nd, 24th, 26th, 28th, 32<sup>nd</sup>, and 33rd times. Therefore, the model proposed in this paper is better.

4.3. *The Evaluation of Physical Health by Linear Sequential Extreme Learning Machine Algorithm Matrix.* Physical health evaluation is the computational goal of differential evolution algorithm to improve vague local search strategy, and it is also an effective embodiment of physical health calculation. The evaluation results of physical health are shown in Figure 7.

It can be seen from Figure 7 that the changes of 35 physical health evaluation indexes are relatively stable,

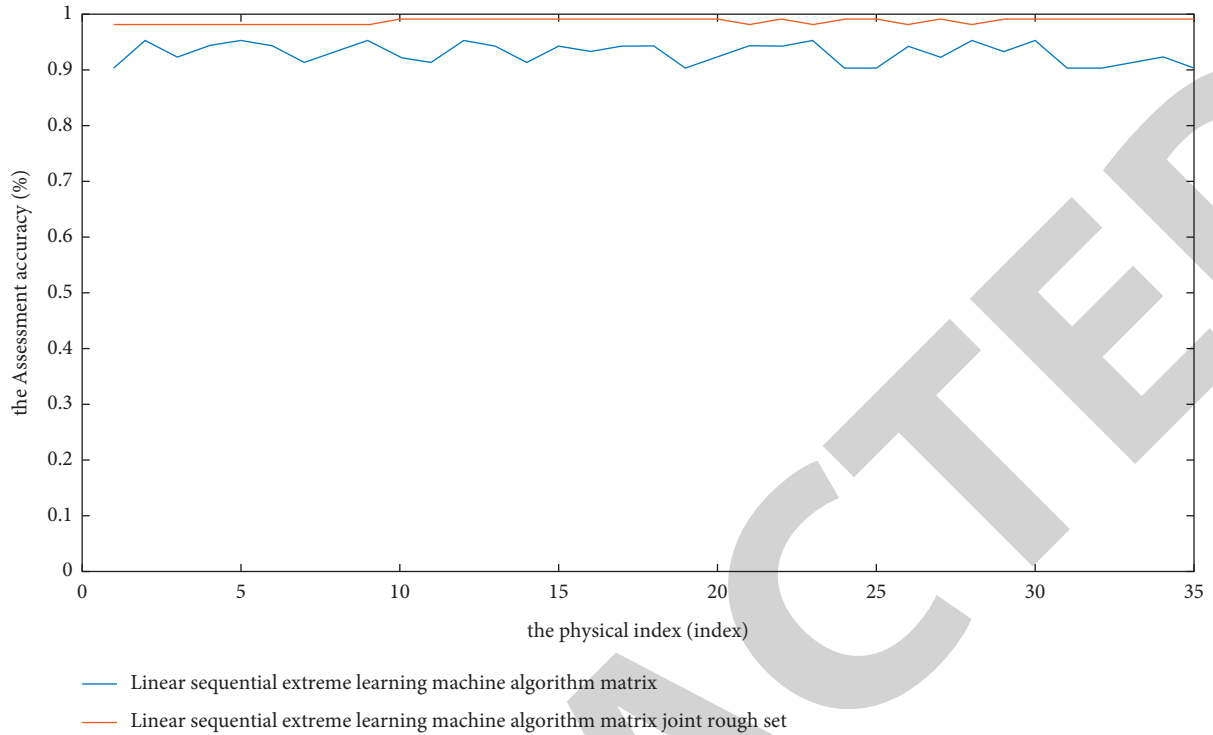


FIGURE 6: The evaluation results of physical health under different algorithms.

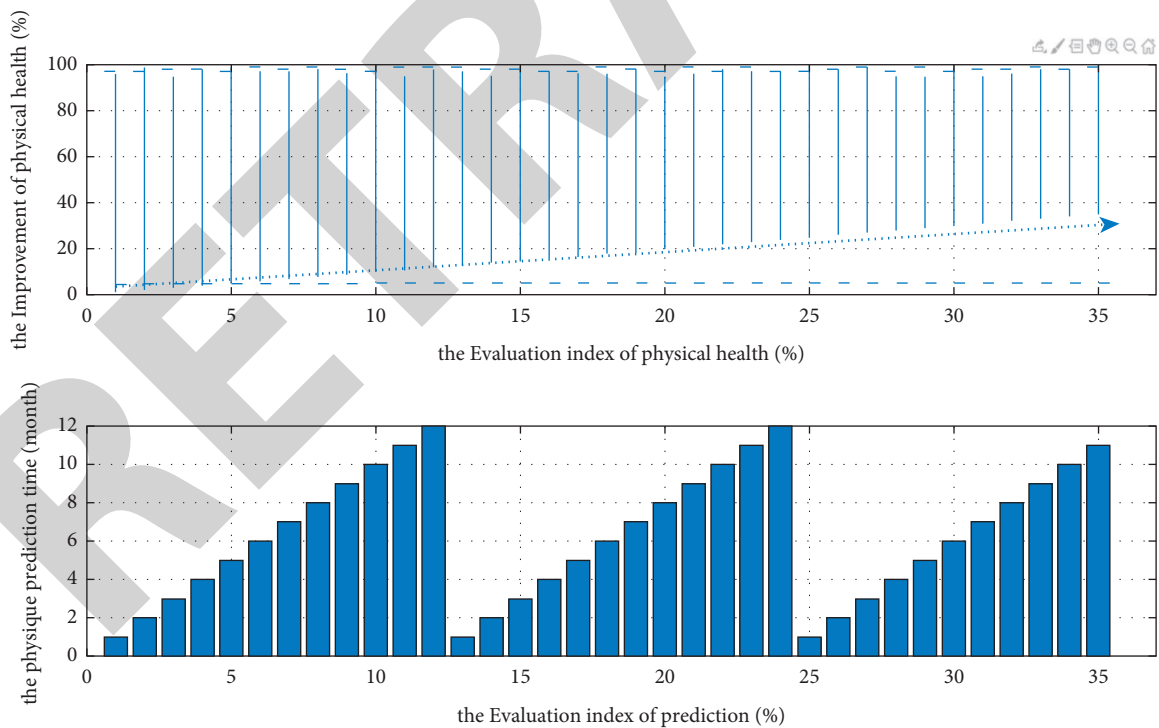


FIGURE 7: The physical health evaluation results.

indicating that the contribution rate of physical health indexes under different algorithms is relatively stable, which verifies the effectiveness of physical health indexes; the results of this study are consistent with relevant domestic studies [23]. At the same time, the evaluation result

of physical health under the linear sequential extreme learning machine algorithm matrix is better, and its change range gradually reduces to 100%. In contrast, the differential evolution algorithm based on vague local search strategy has not changed, and the result is about

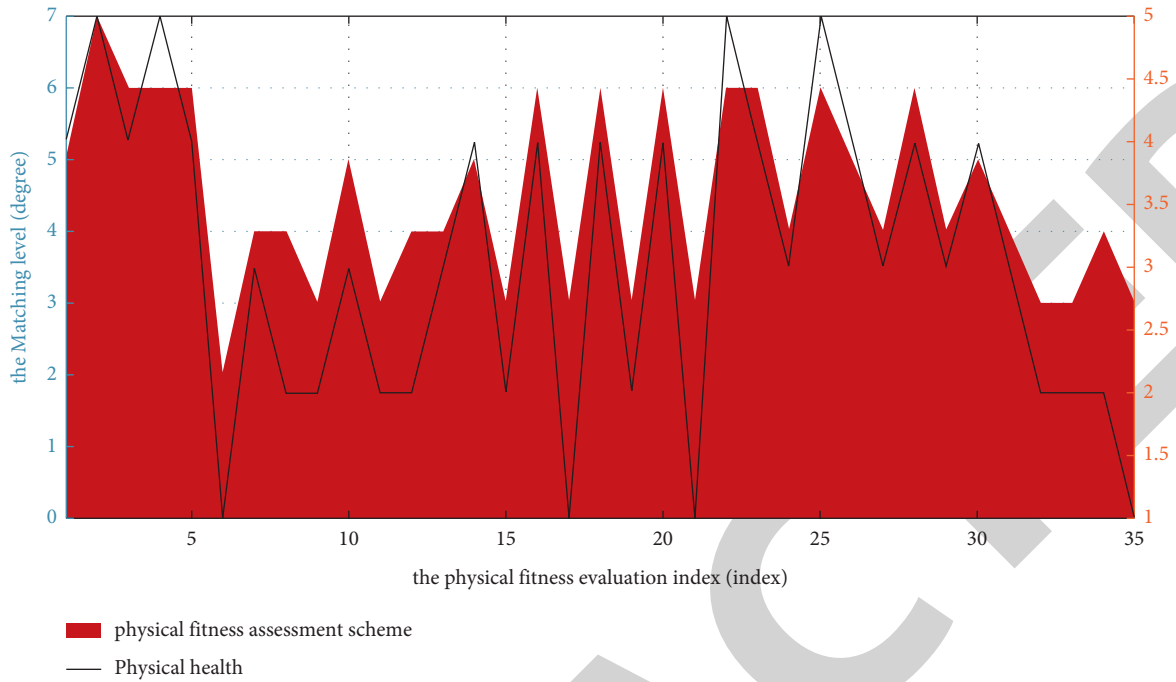


FIGURE 8: The coupling of physical health regimen and physical health.

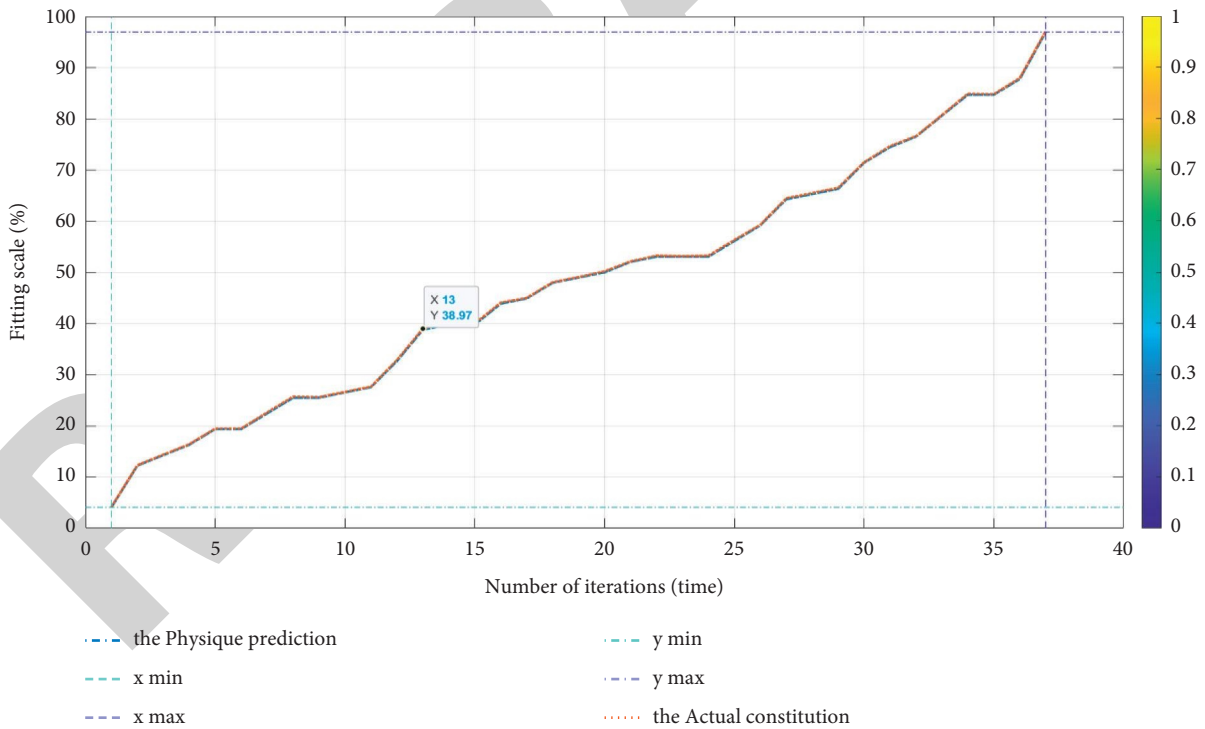


FIGURE 9: The coupling of physical health regimen and different indexes of physical health.

7%. The reason is that the rough set method improves the index aggregation of physical health, provides effective physical fitness assessment schemes for different individuals, and realizes the significant improvement of physical health. Although the differential evolution

algorithm of vague local search strategy also provides effective evaluation scheme for physical health, in the case of randomness, physical fitness assessment scheme presents beneficial and unhelpful changes, resulting in the final improvement of 7%.

**4.4. The Coupling between Physical Fitness Assessment Scheme and Physical Health.** The coupling between physical fitness assessment scheme and physical health is not only the advantage of differential evolution algorithm to improve vague local search strategy, but also the result of two-way evaluation. For better analysis, linker analysis method is selected (uncoupled, general coupling, general, coupling, and comparative coupling, with values of 1–7 respectively). The calculation results are shown in Figure 8.

As can be seen from Figure 8, the coupling between physical fitness assessment scheme and physical health shows an upward trend and approaches the value of 5, indicating that the coupling between the two is better. The results of this study are consistent with relevant domestic studies [24]. Relatively speaking, the linear sequential extreme learning machine algorithm matrix has better coupling, which is significantly better than the differential evolution algorithm with vague local search strategy. The above results also show that rough set can greatly improve the coupling of linear sequential extreme learning machine algorithm matrix and realize the two-way evaluation of physical fitness assessment scheme and physical health. In order to further verify the coupling results, 35 specific indicators are analyzed, and the results are shown in Figure 9.

It can be seen from Figure 9 that the coupling of 35 indicators is better. The results of this study are consistent with relevant domestic studies. Among them, the linear sequential extreme learning machine algorithm matrix has better coupling [Note: 24 red indicators], and the coupling under the two algorithms meets the maximum and minimum constraints of penalty value.

## 5. Conclusion

At present, the research on the impact of physical health on subdivision function has the problem of single evaluation [25], which cannot realize the dual evaluation of physical fitness assessment scheme and the improvement of physical health, so that physical fitness assessment scheme cannot effectively improve physical health. Therefore, how to effectively use physical fitness assessment scheme to improve people's physical health and realize national fitness is an urgent problem to be solved. Based on this background, this paper improves the differential evolution algorithm of vague local search strategy and integrates the rough method set to realize the dual evaluation of physical fitness assessment scheme and physical health. In this paper, firstly, rough set is used to calculate the eigenvalues of physical fitness assessment scheme and physical health evaluation, then rough set is used to eliminate redundant data, and finally the dual evaluation of physical fitness assessment scheme is done. In order to submit the accuracy of evaluation, the constraint conditions and penalty values of rough set are set. MATLAB simulation results show that the linear sequential extreme value learning machine algorithm matrix combined with rough set has a good effect on physical health evaluation and prediction. The result accuracy is greater than 95% and the fitting degree is 98%, which is better than the result of linear sequential extreme value learning machine algorithm

matrix. Therefore, using linear sequential extreme value learning machine algorithm matrix combined with rough set can realize physical health evaluation and prediction, improving the accuracy of the results. However, the interaction between physical health assessment and prediction is not analyzed in the process of rough set analysis in this paper. This aspect will be emphatically analyzed in future research.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This article was supported by the fundamental research funds for the central universities in Chongqing University (2021CDSKXYTY003). The authors appreciate its support very much.

## References

- [1] P. Marwedel, T. Mitra, M. E. Grimheden, and H. A. Andrade, "Survey on education for cyber-physical systems," *IEEE Design & Test*, vol. 37, no. 6, pp. 56–70, 2020.
- [2] P. Mukucha, "The effects of business process outsourcing on the quality of catering services in tertiary education industry in Zimbabwe," *Cogent Business & Management*, vol. 7, no. 1, pp. 32–34, 2020.
- [3] R. S. Nagoyitsyn, "Interactive technologies in developing student's motivation in physical education and sport," *International Journal of Applied Exercise Physiology*, vol. 9, no. 6, pp. 78–85, 2020.
- [4] T. W. Shen, "Development and evaluation of virtual reality induction electricity prevention education and training tools for construction industry," in *Proceedings of the 7th IEEE International Conference on Consumer Electronics - Taiwan (ICCE-Taiwan)*, vol. 8, no. 7, pp. 6–11, Taoyuan, Taiwan, 28–30 Sept. 2020.
- [5] X. Y. Shi, "A didactic pedagogical approach toward sustainable architectural education through robotic tectonics," *Sustainability*, vol. 12, no. 5, pp. 7–10, 2020.
- [6] M. Tornngren, F. Asplund, and M. Magnusson, "The role of competence networks in the era of cyber-physical systems - promoting knowledge sharing and knowledge exchange," *IEEE Design & Test*, vol. 37, no. 6, pp. 8–15, 2020.
- [7] T. H. Nazifa and K. K. Ramachandran, "Exploring the role of information sharing in supply chain management: a case study," *Journal of System and Management Sciences*, vol. 8, no. 4, pp. 13–37, 2021.
- [8] M. M. Alam and F. Faisal, "Mass-private partnership (PPP) projection Bangladesh: current status and challenges," *Journal of System and Management Sciences*, vol. 8, no. 4, pp. 38–56, 2021.
- [9] G. Chen, M. Gao, Z. Zhang, and S. Li, "Hybridization of chaotic grey wolf optimizer and dragonfly algorithm for short-term hydrothermal scheduling," *Ieee Access*, vol. 8, no. 8, pp. 142996–143020, 2020.

## Research Article

# Carbon Emission Calculation and Influencing Factor Analysis Based on Industrial Big Data in the “Double Carbon” Era

Lu Zhang <sup>1,2</sup>, Yan Yan <sup>1</sup>, Wei Xu <sup>1</sup>, Jun Sun <sup>3</sup>, and Yuanyuan Zhang <sup>4</sup>

<sup>1</sup>School of Management, Shenyang University of Technology, Shenyang, Liaoning Province, China

<sup>2</sup>Journal Editorial Department, Shenyang University of Technology, Shenyang, Liaoning Province, China

<sup>3</sup>Liaoning Urban and Rural Gas Co., Ltd, Liaoning Energy Investment (Group) Co., Ltd, Shenyang, Liaoning Province, China

<sup>4</sup>Dalian Medical University, Dalian, Liaoning Province, China

Correspondence should be addressed to Yuanyuan Zhang; zhangyuan@dmu.edu.cn

Received 8 December 2021; Revised 30 December 2021; Accepted 7 January 2022; Published 28 February 2022

Academic Editor: Daqing Gong

Copyright © 2022 Lu Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The arrival of the “double carbon” era indicates that industrial carbon reduction with high energy consumption and high carbon emission is imperative. From the perspective of carbon emission driving factors, industrial carbon emission is decomposed into five influencing factors: energy intensity, energy structure, industrial structure, economic efficiency, and employee scale. Taking the data of 41 subindustries of industrial industry in Liaoning Province from 2010 to 2019 as the research sample, the carbon emission is calculated. The LMDI model is used to analyze and point out the difference in the driving contribution of carbon emissions of each subindustry. The results show that the total carbon emission of Liaoning province gradually decreases, decreases for the first time in 2014, and gradually turns from flat to upward. Economic efficiency is the only and most important reason to drive the increase of industrial carbon emissions in Liaoning Province, and energy efficiency is the primary factor to curb carbon emissions. High carbon industries play a significant role in promoting the increase of carbon emissions, while the medium and low carbon industries have a better effect on restraining carbon emissions. It provides reference and theoretical basis for the government to adjust the industrial structure, control industrial overcapacity, and realize the “double carbon” goal as soon as possible. It is of great significance for the country to optimize energy layout, ensure energy security, and implement the new energy strategy.

## 1. Introduction

As one of the most important material production departments in the national economy, industry is responsible for processing and assembling natural resources and raw materials, providing raw materials, fuel, and driving force for the industry itself and other sectors of the national economy, as well as providing industrial and industrial consumer goods for the people’s material and cultural life. After the industrial revolution, with the large-scale development and use of fossil fuels such as coal, oil, and natural gas, the concentration of carbon dioxide in the Earth’s atmosphere increased significantly, and the global carbon emission reached more than 34 billion tons in 2019. NASA observation data show that compared with the end of the 19th

century, with the increase of global carbon emissions, the current global average temperature has also increased by more than 1.2°C, causing a series of extreme climate problems. More scientists predict that if this trend cannot be curbed, the glaciers in Antarctica will melt at the rate of 250–300 billion tons a year until they disappear completely, which will lead to a rise of at least 56 meters in the global sea level, which is a disaster for mankind. \* \* \* Countries must take decisive steps. China will increase its national independent contribution, adopt more effective policies and measures, strive to reach the peak of carbon dioxide emissions by 2030, and strive to achieve carbon neutrality by 2060. So far, the “double carbon” era has officially arrived, and carbon neutralization and carbon peak have also become the consensus of global green and sustainable



development, marking the largest new industrial revolution since human society entered the industrial revolution. In the “double carbon” era, it is of great significance for the country to further adapt to the global trend and tide by measuring and decoupling the carbon intensity of various industries in order to better achieve the double carbon goal and point out the direction for the industry as the fundamental guarantee of national economic independence, political independence, and national defense modernization. According to statistics, the carbon emissions of the world’s major carbon-emitting countries will reach 34.17 billion tons in 2019; among them, China’s carbon emissions accounted for 28.76, the United States accounted for 14.53, and India accounted for 7.26, ranking among the top three; Russia, Japan, and Germany rank 4th, 5th, and 6th, respectively, with carbon emissions of 1.533 billion tons, 1.123 billion tons, and 684 million tons, respectively [1]. Among China’s domestic provincial carbon emission intensity in 2019, Xinjiang, Shanxi, Inner Mongolia, and Ningxia have higher carbon emission intensity [2].

## 2. Literature Review

From the global perspective: Based on the data of 25 upper-middle-income countries from 1985 to 2014, Mujtaba et al. study the growth, correlation, and causality of kinked index among economic growth, energy consumption, population, trade openness, and carbon dioxide emissions. The results show that economic growth is negatively correlated with CO<sub>2</sub> emissions and trade openness, respectively, while energy consumption is positively correlated with CO<sub>2</sub> emissions and population. The determinants of carbon dioxide emissions are directly proportional to the country’s CO<sub>2</sub> emissions. There is a two-way causal relationship between population and economic growth, and trade openness and economic growth [3]. Georgia proposed that the Royal Court had revealed that in order to achieve net-zero carbon emissions, a series of measures would be taken, including banning some production materials and making 75% of food menus free of meat [4]. André and Valenciano-Salazar considered the environmental management system standards, the AHP method was used to analyze 24 companies and institutions in Costa Rica, and the results showed that the score of carbon neutrality (CN) is higher than the environmental management system standards. Therefore, they suggested that the adoption of environmental certification by organizations should be treated as a multistandard problem considering environmental sustainability, economy, and strategy [5].

From the industries perspective: Demetriou et al. considered that energy advantages and short driving distance make Cyprus a test field for an all-electric vehicle fleet, and they gave four different scenarios of minimum cost (LCSc), business as usual (BAU), carbon capture and storage (CCSc), and renewable energy (RESc) in parallel, deduced the hourly profile of weekly power production and consumption in winter and summer, and predicted that in 2050, RESc will have the highest power generation cost (0.115 euro/kWh) and the highest power loss rate (40%) and need 5600 mwh of domestic battery energy storage, and the current battery capacity in Europe is 3400 mwh [6]. Anonymous pointed out

that the road map of the California Nevada Cement Association (CNCA) showed that by 2045, the three ways for the cement industry to realize net-zero carbon operation were to reduce production process emissions, reduce combustion emissions through fuel switching, and increase distributed power generation [7]. Srivastava et al. discussed the properties of available wastes and different strategies for decomposition or hydrolysis, and efficient microbial systems are emphasized. Some representative examples of biomass sources are used for biomass energy production by providing key information, as well as bioenergy production and environmental problems of plant organics. They also conducted efficient microbial and chemical process research on advanced biofuels extracted from organic matter, so as to promote the use of plant biomass to produce biofuels [8].

Comparing the literature above, the highlight of this paper is that we measure the carbon dioxide emissions of Liaoning industrial segments in different years and analyze the reasons for this phenomenon. At the same time, based on the LMDI decomposition method, we find out the real driving factors behind the growth of carbon dioxide emissions from industrial industries in Liaoning Province and give targeted suggestions from the macro and micro levels, so as to provide a method for the country to conduct horizontal comparison among provinces and vertical comparison among subdivided industries, which is of certain significance to achieve the goal of “double carbon” with high quality.

## 3. Research Method

*3.1. Calculation Example of Carbon Strength.* Carbon dioxide emissions can be divided into natural emissions and artificial emissions. Natural emission refers to the emission of carbon dioxide into the atmosphere from soil, ocean, forest, and other ecosystems; artificial emissions refer to carbon dioxide emissions caused by human activities, mainly from fossil energy consumption and biomass fuel combustion. There are three main research methods on carbon dioxide emissions by scholars at home and abroad, namely, emission coefficient method, material balance algorithm, and measurement method. Because the research object of this paper is industrial carbon emissions, it is mainly aimed at fossil energy, and the emission coefficient method is used to calculate. The calculation coefficients of various energy sources are shown in Table 1. The calculation formula is based on the reference formula proposed by the United Nations Intergovernmental Panel on climate change (IPCC) in 2006 (this part of the 2019 inventory guide has not changed) [10], combined with the availability and acceptability of variable data and the specific characteristics of national industries, and the calculation way is as follows:

$$\begin{aligned} \text{CO}_2 &= \sum_{i=1}^N (\text{CO}_2)_i \\ &= \sum_{i=1}^N E_i \times \text{NCV}_i \times \text{CC}_i \times \text{COF}_i \times \frac{44}{12} \end{aligned} \quad (1)$$

TABLE 1: Calculation correlation coefficient of various energy sources [9].

Energy name	Conversion coefficient of standard coal	Average low calorific value	Carbon content per unit calorific value (ton carbon/TJ)	Carbon oxidation rate	Carbon dioxide emission coefficient
Raw coal	0.7143 kgce/kg	20908 kJ/kg	26.37	0.94	1.9003 kg-CO <sub>2</sub> /kg
Gasoline	1.4714 kgce/kg	43070 kJ/kg	18.9	0.98	2.9251 kg-CO <sub>2</sub> /kg
Diesel oil	1.4571 kgce/kg	42652 kJ/kg	20.2	0.98	3.0959 kg-CO <sub>2</sub> /kg
Fuel oil	1.4286 kgce/kg	41816 kJ/kg	21.1	0.98	3.1705 kg-CO <sub>2</sub> /kg
Natural gas	1.3300 kgce/m <sup>3</sup>	38931 kJ/m <sup>3</sup>	15.3	0.99	2.1622 kg-CO <sub>2</sub> /m <sup>3</sup>

where  $I$  represents the type of energy;  $E$  is energy consumption; NCV is the average low calorific value; CC is carbon content; COF is carbon oxidation factor, expressed as a percentage of IPCC default value; 44 and 12 are the molecular weights of CO<sub>2</sub> and C, respectively.

According to China's national standards, the division of industrial industries and the classification of main energy consumption are shown in Figure 1.

Since *China's industrial Statistical Yearbook* will no longer publish a separate total industrial output value of subdivided industries after 2009, *Liaoning Provincial Statistical Yearbook* has an accurate calculation of the data. Considering the availability and accuracy of the data, 41 subindustries of the Liaoning industry from 2010 to 2019 are selected as the research object. The data on the energy consumption of kerosene and coke are short. In order to ensure the accuracy of the calculation, these two data are excluded, so five raw materials of fossil energy raw coal, gasoline, fuel oil, diesel, and natural gas are selected for calculation, and a few missing data are estimated and completed by linear interpolation method. Relevant data are from the 2006 IPCC guidelines for national greenhouse gas inventories, *Liaoning Statistical Yearbook* and *China Energy Statistical Yearbook*. After calculation and sorting, the change trend of total carbon emission and total industrial output value of each subindustry of Liaoning industry from 2010 to 2019 is shown in Figure 2, and the carbon emission is shown in Table 2 (considering comprehensive application, international practice is adopted).

**3.2. Results' Analysis.** In 2019, the industrial economy of Liaoning Province continued to optimize and upgrade the industrial structure along with the government's cultivation and growth of new industries and upgrading of traditional industries. However, affected by the escalation of Sino-US trade friction, the larger decline in the price of industrial products, the lack of domestic demand, and other factors, the industrial economy of Liaoning Province as a whole presents a situation of great difficulty in production and operation, slow pace, and falling profits. According to the calculation of carbon emissions of 41 industrial subindustries in Liaoning Province in recent ten years, the total carbon emissions of electric power, thermal power production, and supply industry and ferrous metal smelting and rolling processing industry are large, which are 162.0196 million T and 75.6808 million T, respectively,

ranking first and second in 41 subindustries, accounting for 68.7% of the total carbon emissions of the whole industrial industry.

Compared with 2010, with the continuous progress of industrial production means, the downward trend of carbon emissions of most industries is more obvious. However, some industries still have the phenomenon of increasing carbon emissions instead of reducing carbon emissions. The largest increase is in the oil, coal, and other fuel processing industry, reaching 218.9%. The comprehensive utilization of waste resources followed, with an increase of 199.57%, and the carbon emission increased from 46700 t in 2010 to 139900 t in 2019. On the contrary, most equipment manufacturing industries have performed well, and the goal of energy conservation and emission reduction has been basically achieved through the joint action of various factors such as industrial upgrading, technological progress, and talent drive.

In 2019, the added value of industries above the designated size increased by 6.7% over 2018. Among them, the added value of the mining industry increased by 2.0% over the previous year, the added value of the manufacturing industry increased by 7.6%, and the added value of the high-tech manufacturing industry increased by 18.7%. The added value of power, heat, gas, and water production and supply also increased by 5.1% over last year.

In terms of specific industries, the added value of 25 of the 41 major industries maintained year-on-year growth, with a growth rate of 61.0%, and one industry was flat. Agricultural and sideline food processing industry increased by 7.1%, food manufacturing industry and textile industry decreased by 9.3% and 14.0%, respectively, the added value of petroleum, coal, and other fuel processing industry increased by 14.9%, the added value of chemical raw materials and chemical products manufacturing industry increased by 11.1%, pharmaceutical manufacturing industry increased by 9.1%, ferrous metal smelting and rolling processing industry, and nonferrous metal smelting and rolling processing industry increased by 5.0% and 3.3%, respectively, the nonmetallic mineral products industry increased by 2.7%, the general equipment manufacturing industry increased by 5.0%, and the special equipment manufacturing industry decreased by 2.1%; in terms of transportation, the automobile manufacturing industry increased by 2.5%, and the railway, ship, aerospace, and other transportation equipment manufacturing industry increased by 4.5%.

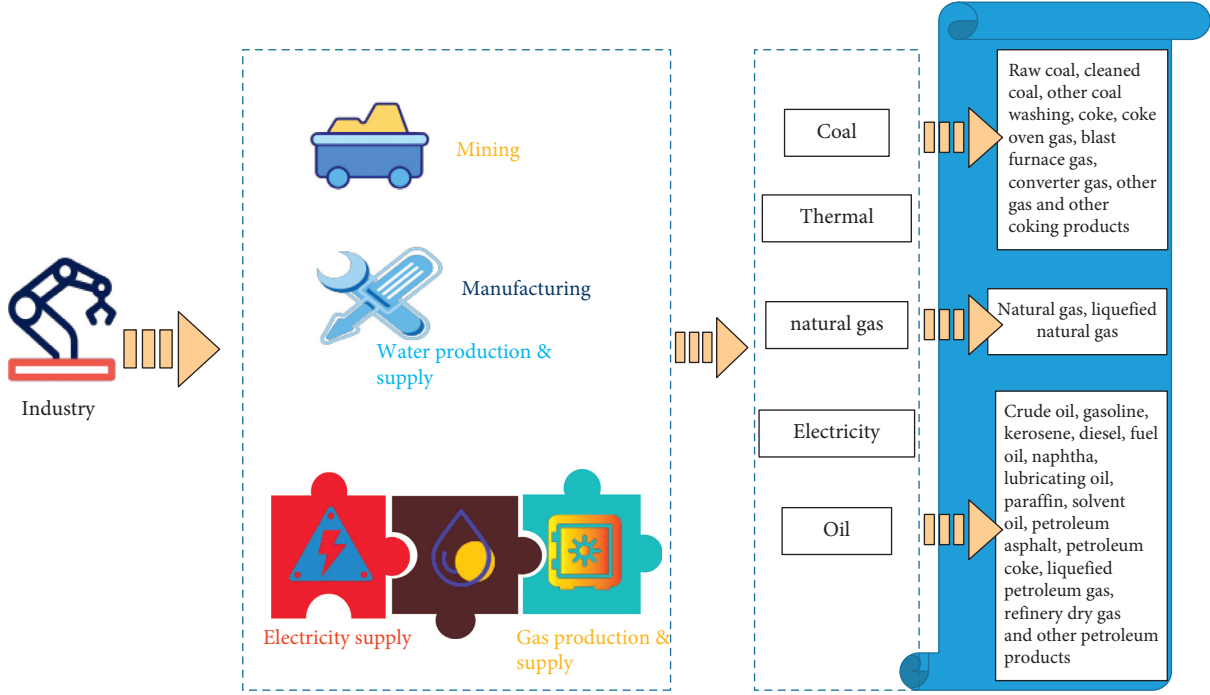


FIGURE 1: Industrial industry division and main energy consumption classification under national standards.

3.3. *Analysis of Influencing Factors.* Decomposing the core influencing factors of carbon emissions and studying the contribution of various factors to carbon emissions can not only provide theoretical reference for the government to formulate policies for phased economic and social green development but also provide a solid foundation for the country to achieve double carbon goals, maintain the bottom line of energy security, and lead high-quality development of the country. At present, the most commonly used methods for calculating carbon emissions are the following:

- (1) IPAT model: a quantitative model used to analyze the relationship between human activities and the environment. With the rapid growth of population and the progress of science and technology, people's pursuit of material life and the exploitation and utilization of resources have brought great pressure to the Earth's resources and environment. In other words, resources and environment are limited, but the human desire for material pursuit is unlimited. The model represents the pressure of human beings on the environment as the product of three factors, namely, the number of people, the degree of abundance, and the progress of science and technology, that is,  $I = P (\text{Pop}) \times A (\text{Aff}) \times T (\text{Tech})$ . Although the method is clear and simple in the calculation, it is not considered that environmental, resource, and social development are not linear in practice.
- (2) EKC: the environmental Kuznets curve, which was put forward by American economist Simon Smith Kuznets in 1955, mainly outlines the changes in the distribution of income along with the economic development process, also known as the inverted U

curve. The formula is expressed as  $E = \beta_0 + \beta_1 Y + \beta_2 Y^2 + \mu$ . Among them,  $E$  is the environmental pressure of the country or region, and  $Y$  is the economic output of the country or region. This method holds that when the economic development reaches a critical point, environmental pollution will tend to be stable no matter how the per capita income increases.

- (3) STIRPAT model: this method extends the IPAT model, adds randomness, and breaks through the limitation of unit single line assumption, and the calculation formula is expressed as  $I_i = \delta_0 P_i^{\delta_1} A_i^{\delta_2} T_i^{\delta_3} e_i$ .
- (4) LMDI model: compared with the first three methods, this method does not need to calculate with the help of an input-output table [11]. It is relatively simple to use, involves the nature of time series, and has more consistent aggregation, zero residual, and independent path [12]. Since the LMDI decomposition method has the advantages of complete decomposition, aggregation consistency, and path independence, this method is adopted in this paper.

The carbon emissions are decomposed in (1), as shown in the following:

$$\begin{aligned}
 C &= \sum_{i=1}^I \sum_{j=1}^J C_{ij} \\
 &= \sum_{i=1}^I \sum_{j=1}^J \frac{C_{ij}}{E_{ij}} \times \frac{E_{ij}}{Y_j} \times \frac{Y_j}{Y} \times \frac{Y}{P} \times P,
 \end{aligned} \tag{2}$$

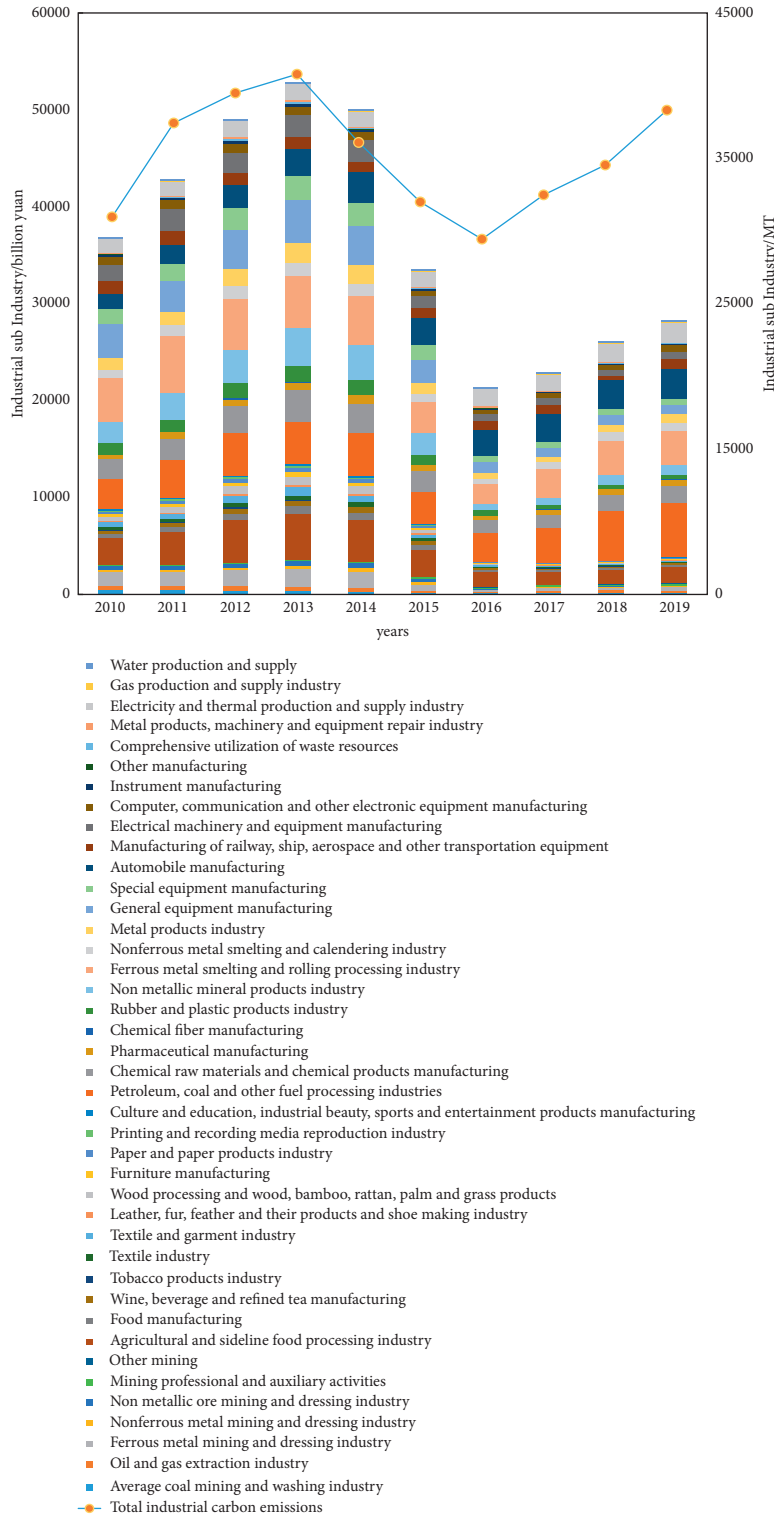


FIGURE 2: Change trend of total carbon emission and output value of various subindustries of Liaoning industry from 2010 to 2019.

where  $C_{ij}$  represents the carbon emission of energy type  $i$  in  $j$  industry;  $E_{ij}$  represents the energy  $i$  consumption of  $j$  industry;  $E_j$  represents the total energy consumption of  $j$  industry;  $Y_j$  represents the total output value of the industrial  $j$  industry;  $Y$  represents the total industrial output value.  $ED_{ij} = C_{ij}/E_{ij}$  represents the carbon emission intensity of

energy  $i$  in industrial  $j$  industry;  $EM_{ij} = E_{ij}/E_j$  represents the energy structure of energy  $i$  in industrial  $j$  industry;  $ET_j = E_j/Y_j$  represents the energy efficiency of energy  $i$  in industrial  $j$  industry;  $ES_j = Y_j/Y$  represents the industrial structure of the industrial  $j$  industry;  $EE = Y/P$  represents the economic efficiency of energy  $i$  in industrial  $j$  industry;  $P$  represents the

TABLE 2: Total carbon emissions of various subindustries of Liaoning industry from 2010 to 2019.

Industrial subindustry	2010 (MT)	2011 (MT)	2012 (MT)	2013 (MT)	2014 (MT)	2015 (MT)	2016 (MT)	2017 (MT)	2018 (MT)	2019 (MT)	Mean value (MT)
Average coal mining and washing industry	2819.37	3009.68	4618.05	2935.12	2617.01	2385.85	2312.94	2794.61	2837.36	2231.74	2856.17
Oil and gas extraction industry	445.39	567.75	452.97	481.03	303.20	299.39	278.40	312.68	328.80	286.56	375.62
Ferrous metal mining and dressing industry	130.28	188.98	156.51	156.45	127.46	121.11	54.91	71.83	98.30	101.52	120.74
Nonferrous metal mining and dressing industry	130.91	186.10	184.23	130.90	169.37	116.38	72.35	97.97	85.41	166.72	134.03
Nonmetallic ore mining and dressing industry	120.99	140.73	210.81	264.13	146.27	116.96	81.05	58.38	54.91	67.34	126.16
Mining professional and auxiliary activities	0.00	0.00	199.28	190.83	158.56	74.53	52.83	52.70	50.16	54.05	83.29
Other mining	0.56	0.56	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.11
Agricultural and sideline food processing industry	296.42	324.90	312.05	347.33	397.31	271.59	257.07	244.71	220.10	216.53	288.80
Food manufacturing	68.24	63.01	66.91	68.65	47.95	38.59	30.82	23.59	23.74	18.40	44.99
Wine, beverage, and refined tea manufacturing	96.84	94.40	95.12	84.06	70.58	55.75	57.84	41.28	35.46	29.47	66.08
Tobacco products industry	1.36	1.60	1.93	1.94	1.85	1.94	1.11	1.30	1.36	1.37	1.58
Textile industry	58.07	52.62	45.09	47.42	39.46	26.79	27.78	24.33	22.06	19.02	36.26
Textile and garment industry	66.31	64.22	54.49	50.43	42.62	20.67	21.87	12.47	5.97	5.47	34.45
Leather, Fur, feather, and their products and shoe-making industry	5.72	4.62	5.58	4.71	3.55	2.31	2.67	0.29	0.12	0.12	2.97
Wood processing and wood, bamboo, rattan, palm, and grass products	58.00	56.63	56.08	44.38	29.07	16.42	6.94	2.02	1.49	1.16	27.22
Furniture manufacturing	31.45	24.88	21.13	17.75	12.84	7.92	4.83	1.43	0.73	0.56	12.35
Paper and paper products industry	151.33	126.68	147.53	146.44	67.75	51.79	41.43	113.47	106.10	122.50	107.50
Printing and recording media reproduction industry	10.37	12.05	10.34	10.29	8.66	3.42	3.60	0.58	0.53	0.42	6.03
Culture and education, industrial beauty, sports, and entertainment products manufacturing	3.70	6.25	16.97	10.52	8.21	4.79	3.89	0.50	0.41	0.38	5.56

TABLE 2: Continued.

Industrial subindustry	2010 (MT)	2011 (MT)	2012 (MT)	2013 (MT)	2014 (MT)	2015 (MT)	2016 (MT)	2017 (MT)	2018 (MT)	2019 (MT)	Mean value (MT)
Petroleum, coal, and other fuel processing industries	1198.60	1396.47	1853.09	2101.50	2060.72	2110.48	2070.27	2353.59	2423.89	3822.34	2139.10
Chemical raw materials and chemical products manufacturing	1378.29	1487.34	1308.66	1342.52	1349.04	1048.84	835.01	758.02	487.04	516.22	1051.10
Pharmaceutical manufacturing	141.19	170.52	151.20	140.00	116.70	96.74	83.16	75.93	66.75	60.47	110.27
Chemical fiber manufacturing	97.33	89.01	86.30	81.48	67.78	77.62	74.84	63.52	10.27	9.87	65.80
Rubber and plastic products industry	213.13	232.42	215.56	212.53	141.94	118.78	61.42	31.59	27.80	27.11	128.23
Nonmetallic mineral products industry	2188.78	2270.55	3017.66	2876.19	2420.39	2015.23	1558.55	1622.11	1585.57	1992.49	2154.75
Ferrous metal smelting and rolling processing industry	7243.14	8029.31	7614.93	7745.26	7747.74	7501.76	7375.67	7235.43	7433.70	7753.88	7568.08
Nonferrous metal smelting and calendering industry	241.14	220.28	172.15	187.26	122.90	113.72	83.09	106.58	129.75	117.92	149.48
Metal products industry	119.49	136.20	145.51	172.42	78.89	50.72	42.61	28.36	73.75	82.56	93.05
General equipment manufacturing	318.81	363.12	316.55	360.04	207.16	192.36	44.86	17.03	14.70	12.93	184.76
Special equipment manufacturing	118.42	148.77	203.50	222.42	92.39	75.28	103.11	48.83	6.73	5.87	102.53
Automobile manufacturing	122.08	166.10	127.67	171.93	75.34	56.74	52.72	37.11	31.20	29.74	87.06
Manufacturing of railway, ship, aerospace, and other transportation equipment	104.98	84.32	127.74	210.81	28.94	19.82	22.13	21.70	53.33	11.91	68.57
Electrical machinery and equipment manufacturing	119.05	135.08	151.85	131.99	94.44	66.41	43.55	18.73	11.42	7.47	78.00
Computer, communication, and other electronic equipment manufacturing	13.62	13.19	10.69	16.13	7.22	5.10	4.67	2.95	2.74	2.07	7.84
Instrument manufacturing	13.40	14.92	14.81	19.91	4.27	2.64	3.27	1.34	1.03	0.98	7.66
Other manufacturing	26.86	41.05	18.94	10.88	2.38	1.61	2.09	0.64	0.12	0.14	10.47



TABLE 2: Continued.

Industrial subindustry	2010 (MT)	2011 (MT)	2012 (MT)	2013 (MT)	2014 (MT)	2015 (MT)	2016 (MT)	2017 (MT)	2018 (MT)	2019 (MT)	Mean value (MT)
Comprehensive utilization of waste resources	4.67	10.73	18.56	17.17	16.08	6.63	11.08	8.35	10.09	13.99	11.74
Metal products, machinery, and equipment repair industry	0.00	0.00	22.53	7.60	3.27	2.10	2.44	15.14	3.78	2.82	5.97
Electricity and thermal production and supply industry	14565.17	15651.38	14170.87	16059.87	16067.43	15954.85	16217.73	17062.14	18065.17	18204.99	16201.96
Gas production and supply industry	4.59	19.37	122.30	23.60	59.06	54.18	41.09	43.87	14.24	12.39	39.47
Water production and supply industry	9.19	13.97	13.98	14.51	7.37	8.44	8.75	5.52	3.79	2.64	8.82
Total industrial carbon emissions	32737.24	35619.76	36540.12	37118.4	35023.17	33196.25	32054.44	33412.62	34329.87	36014.13	

number of employees. The decomposition formula of each carbon emission factor is calculated as shown in (3)–(8).

$$D_{ED} = \sum_{i=1}^I \sum_{j=1}^J \omega_{ij} \ln \frac{ED_{ij}^{t+1}}{ED_{ij}^t}, \quad (3)$$

$$D_{EM} = \sum_{i=1}^I \sum_{j=1}^J \omega_{ij} \ln \frac{EM_{ij}^{t+1}}{EM_{ij}^t}, \quad (4)$$

$$D_{ET} = \sum_{i=1}^I \sum_{j=1}^J \omega_{ij} \ln \frac{ET_j^{t+1}}{ET_j^t}, \quad (5)$$

$$D_{ES} = \sum_{i=1}^I \sum_{j=1}^J \omega_{ij} \ln \frac{ES_j^{t+1}}{ES_j^t}, \quad (6)$$

$$D_{EE} = \sum_{i=1}^I \sum_{j=1}^J \omega_{ij} \ln \frac{EE_i^{t+1}}{EE_i^t}, \quad (7)$$

$$D_{EP} = \sum_{i=1}^I \sum_{j=1}^J \omega_{ij} \ln \frac{P_j^{t+1}}{P_j^t}. \quad (8)$$

Among them,  $\omega_{ij}$  is the weight of energy type  $I$  in  $J$  industry, and the calculation formula is shown in the following:

$$\begin{aligned} \omega_{ij} &= L(C_{ij}^{t+1}, C_{ij}^t) \\ &= \frac{C_{ij}^{t+1} - C_{ij}^t}{\ln C_{ij}^{t+1} - \ln C_{ij}^t}. \end{aligned} \quad (9)$$

In view of using the two forms of “addition” and “multiplication” in the LMDI decomposition model, the final results are the same, and the addition is more intuitive

and clear than “multiplication”, so “addition” is selected for factor decomposition in this paper, and the calculation formula is shown in the following:

$$IC^{t+1} - CI^t = D_{ED} + D_{EM} + D_{ET} + D_{ES} + D_{EE} + D_{EP}. \quad (10)$$

**3.4. Case Analysis.** As a major industrial province, Liaoning has always undertaken the important task of providing technical equipment for various industries of China’s national economy and national defense construction and has played an indispensable role in the whole process of industrialization. In 2020, faced with the impact of COVID-19 and the severe external environment of the world, Liaoning always adhered to the working keynote of steady progress. The industrial added value of above scale industries increased by 1.8% compared with that in 2019, of which the mining industry decreased by 0.6% compared with the same period last year, and the manufacturing industry grew by 1.8% over the same period last year. The electricity, water, gas, and aquatic industries and supply also increased by 3.6% over the same period last year.

Considering the national development of the green economy and the realization of the “double carbon” goal, Liaoning needs to balance economic growth and carbon emission, so take Liaoning as an example for LMDI decomposition analysis. In view of space, 41 industrial industries in Liaoning Province are numbered, as shown in Table 3. According to the carbon emissions of each sub-industry, it is divided into high, medium, and low carbon industries, as shown in Table 4.

According to the LMDI model, 2010–2019 is selected as the research period to decompose the total industrial carbon emission in Liaoning Province and calculate the contribution degree and direction of five influencing factors such as energy intensity, energy structure, industrial structure,

TABLE 3: Industrial code of Liaoning Province.

Number	Industry name
1	Average coal mining and washing industry
2	Oil and gas extraction industry
3	Ferrous metal mining and dressing industry
4	Nonferrous metal mining and dressing industry
5	Nonmetallic ore mining and dressing industry
6	Mining professional and auxiliary activities
7	Other mining
8	Agricultural and sideline food processing industry
9	Food manufacturing
10	Wine, beverage, and refined tea manufacturing
11	Tobacco products industry
12	Textile industry
13	Textile and garment industry
14	Leather, fur, feather, and their products and shoe-making industry
15	Wood processing and wood, bamboo, rattan, palm, and grass products
16	Furniture manufacturing
17	Paper and paper products industry
18	Printing and recording media reproduction industry
19	Culture and education, industrial beauty, sports, and entertainment products manufacturing
20	Petroleum, coal, and other fuel processing industries
21	Chemical raw materials and chemical products manufacturing
22	Pharmaceutical manufacturing
23	Chemical fiber manufacturing
24	Rubber and plastic products industry
25	Nonmetallic mineral products industry
26	Ferrous metal smelting and rolling processing industry
27	Nonferrous metal smelting and calendering industry
28	Metal products industry
29	General equipment manufacturing
30	Special equipment manufacturing
31	Automobile manufacturing
32	Manufacturing of railway, ship, aerospace, and other transportation equipment
33	Electrical machinery and equipment manufacturing
34	Computer, communication, and other electronic equipment manufacturing
35	Instrument manufacturing
36	Other manufacturing
37	Comprehensive utilization of waste resources
38	Metal products, machinery, and equipment repair industry
39	Electricity and thermal production and supply industry
40	Gas production and supply industry
41	Water production and supply

TABLE 4: Carbon emission classification of industrial subindustries in Liaoning.

Carbon emission classification	Subindustry name and number
High carbon	1, 4, 6, 7, 20, 37, 39
Medium carbon	2, 3, 5, 8, 9, 10, 11, 12, 17, 21, 22, 23, 24, 25, 26, 27, 28, 31, 32, 34
Low carbon	13, 14, 15, 16, 18, 19, 29, 30, 33, 35, 36, 38, 40, 41

economic efficiency, and population effect to industrial carbon emission. The results are shown in Table 5.

In terms of energy intensity, from 2010 to 2019, the industrial energy intensity of Liaoning Province generally showed a downward trend, indicating that the industrial production technology of Liaoning Province has been improved to a certain extent. In terms of energy structure, from 2010 to 2019, the change of energy structure leads to the phenomenon of positive and negative alternation of

industrial carbon emission year by year, with large fluctuation. The final cumulative carbon emission is  $-65.5186$  million T, and compared with the other four influencing factors, the impact effect is the weakest, and the contribution rate is only 0.75%. In terms of industrial structure, most of the carbon emissions during the study period were negative, the cumulative carbon emissions decreased by 781259.98 million T, and the contribution rate to carbon emissions was 89.35%, indicating that it is the most important inhibitory

TABLE 5: Effect decomposition results of influencing factors of industrial carbon emission in Liaoning Province from 2010 to 2019.

Year	$D_{EM}$	$D_{ET}$	$D_{ES}$	$D_{EE}$	$D_{EP}$
2010-2011	-2182.06	-84889.77	-382022.83	1325.17	3842.70
2011-2012	2014.54	6223.20	15690.64	5010.07	-144.43
2012-2013	-1886.00	-6095.04	35906.64	823.19	1967.99
2013-2014	2124.50	1371.04	-24989.56	586.65	-2549.06
2014-2015	-795.55	2805.60	-54870.20	-10415.13	-3304.94
2015-2016	-2871.30	7341.45	-144101.15	-10683.17	-4059.59
2016-2017	2946.92	-1947.07	248635.79	5556.19	-3144.13
2017-2018	-3149.61	-3009.94	-425147.36	5739.87	-1425.34
2018-2019	-2753.30	-275.44	-50361.95	3531.02	-737.02
Cumulative value	-6551.86	-78475.97	-781259.98	1473.86	-9553.82

factor of carbon emissions and plays a vital role in reducing carbon emissions in Liaoning Province. Economic efficiency is closely related to the total industrial output value, reflecting the progress of the economic level and technical means of the whole industry. As can be seen from Table 5, the cumulative increase of carbon emissions caused by economic efficiency has reached 1473.86 million T, which is the only driving factor leading to the increase of carbon emissions. The population effect reflects the attribute of social evaluation. During 2010–2019, the contribution rate to carbon emission reached 1.09%, second only to economic efficiency.

From 2010 to 2019, the cumulative carbon emission of industrial high carbon industries in Liaoning Province reached 39.15194 million T. The industry of comprehensive utilization of waste resources has the highest growth rate, accounting for 30.67% of the cumulative carbon emissions. The industry of oil, coal, and other fuel processing ranked second, with cumulative carbon emissions accounting for 29.65%. Among the influencing factors, most of the energy structure has a positive role in promoting, and only has a certain inhibitory effect on mining professionals and auxiliary activities; energy efficiency has the most obvious inhibitory effect on electricity, thermal production, and supply industry. The driving force of industrial structure on mining professional and auxiliary activities is relatively strong; economic efficiency and the size of employees have contributed a certain inhibitory force to the increase of carbon emissions in industrial high-carbon industries, but the emission reduction effect is also very weak in some industries, as shown in Table 6.

In the ten years from 2010 to 2019, the carbon emission of Liaoning's industrial carbon industry reached -1640377600 t. Carbon emissions from all industrial sectors have played a negative inhibitory role. From the perspective of influencing factors, the inhibition of energy structure is most significant, followed by energy efficiency indicators. The index of industrial structure has played a certain role in promoting the increment of industrial carbon emissions, as shown in Table 7.

During the study period, the cumulative carbon emissions of Liaoning's industrial low carbon industry reached -413557.2 million T. The driving force of promoting carbon emissions is strong in instrumentation manufacturing, gas production and supply, textile and clothing, and apparel industries. From the point of view of influencing factors,

energy structure only has positive effects on gas production and supply industry and has played an inhibitory role in other low carbon industries. Energy efficiency has only played a catalytic role in the manufacturing of culture, education, industry, sports, and entertainment products, and the emission reduction effect of the remaining low carbon industries is outstanding. The three factors, including industrial structure, economic sub-Olympic Green, and the scale of employees, played a positive role in promoting the increase of carbon emissions in some low-carbon industries, but the impact was weak, as shown in Table 8.

#### 4. Results and Discussion

From the perspective of the whole industrial sector, the total carbon emission of 41 industrial subindustries in Liaoning Province increased from 327372400 t to 360141300 t from 2010 to 2019, with an absolute increase of 1.1 times. From the perspective of evolution trend, the growth rate of total industrial carbon emission in Liaoning province gradually decreased from 2010 to 2013, and the carbon emission decreased for the first time after 2014. The decline rate of total carbon emission in Liaoning province gradually leveled off from 2014 to 2016, and there was an upward trend from 2016 to 2019.

From the analysis results of influencing factors and indicators, economic efficiency is the main and only reason driving the increase of carbon emissions from industrial industries in Liaoning Province; energy efficiency is the primary factor to restrain the carbon emission of industrial industry in Liaoning Province, followed by the industrial structure, while the impact of employee size and energy structure on the carbon emission of the whole industry is relatively weak.

From the perspective of the fine molecule industry, the high carbon industry plays an extremely significant role in promoting the increase of industrial carbon emissions, among which the driving role of comprehensive utilization of waste resources is the most prominent; medium and low carbon industries have good effects on restraining industrial carbon emissions in Liaoning Province. Therefore, the carbon emissions of the industrial sector mainly come from high carbon emission industries, so it is necessary to carry out energy conservation, emission reduction, upgrading, and transformation for subindustries.

TABLE 6: Decomposition of influencing factors of high-carbon industry emission.

Industry number	$D_{EM}$	$D_{ET}$	$D_{ES}$	$D_{EE}$	$D_{EP}$	Cumulative value
1	-1401.61	2278.35	-2808.45	1577.33	2742.41	2388.03
4	424.12	2576.38	-906.95	1120.25	-1592.19	1621.61
6	-27321.17	-7467.2	19408.79	-23655.51	43054.96	4019.87
7	45913.81	27694.5	-35603.48	-22145.29	-12710.51	3149.03
20	4390.79	853.83	4643.59	1670.32	48.94	11607.47
37	4653.54	16149.4	-10502.37	-5013.78	6719.32	12006.11
39	449.55	-13105.63	14880.7	2362.52	-227.32	4359.82

TABLE 7: Decomposition of influencing factors of medium-carbon industry emission.

Industry number	$D_{EM}$	$D_{ET}$	$D_{ES}$	$D_{EE}$	$D_{EP}$	Cumulative value
2	-1954.24	-765.84	258.51	3187.93	-4809.65	-4083.29
3	-776.58	3674.68	-3017.86	-3529.69	-1017.05	-4666.5
5	-1892.08	3859.53	-4617.92	-2630.71	-2928.64	-8209.82
8	-1319.76	-5874.43	5950.63	-837.32	52.89	-2027.99
9	-5411.27	-3464.94	-634.47	-2346.2	-359.91	-12216.79
10	-4923.52	-8808.6	5251.45	-356.88	-2394.72	-11232.27
11	1292.57	-7236.2	-4804.68	-2046.37	-1234.21	-14028.89
12	-4618.68	3567.67	-6738.01	488.16	-6486.22	-13787.08
17	-1395.08	-1907.97	1760.38	-1232.53	-2123.94	-4899.14
21	-3990.99	-9009.1	6436.27	2718.01	-2713.72	-6559.53
22	-3678.45	-12873	10579.15	1154.72	274.17	-4543.41
23	-9047.62	493.57	-8218.17	3439.91	-5764.25	-19096.56
24	-8102.93	4797.24	-1944.47	-3106.02	-811.04	-9167.22
25	-667.85	2268.82	-1740.46	-794.77	-1195.08	-2129.34
26	-132.99	1179.06	-124.64	369.08	-1755.25	-464.74
27	-3354.22	-2214.37	318.9	936.92	-1143.38	-5456.15
28	-1344.72	-173.52	137.18	-498.4	378.39	-1501.07
31	-5946.8	-5804.74	1354.16	1029.88	2064.06	-7303.44
32	-9714.97	-7651.11	-1177.19	-353.38	-434.54	-19331.19
34	-7687.81	-12436.79	6090.77	1456.52	-756.03	-13333.34

TABLE 8: Decomposition of influencing factors of low-carbon industry emission.

Industry number	$D_{EM}$	$D_{ET}$	$D_{ES}$	$D_{EE}$	$D_{EP}$	Cumulative value
13	-9693.74	-10924.28	2769.32	-1739.89	-2120.64	-21709.23
14	-15265.81	-6226.96	-7639.47	-4043.53	-1613.74	-34789.51
15	-15367.17	-9911.17	-4028.89	-4637.31	-2171.66	-36116.2
16	-16212.24	-10174.88	-4553.8	-3468.14	-1361.09	-35770.15
18	-12694.85	-11565.35	321.02	-2099.47	-2496.68	-28535.33
19	-8589.59	9440.41	-17278.95	-6669.27	-1132.25	-24229.65
29	-12602.5	-7342.06	-3784.34	-2731.48	-1214.46	-27674.84
30	-12099.44	-7934.57	-2922.32	-3774.22	-107.65	-26838.2
33	-11154.18	-10105.49	59.26	-2051.8	-1120.54	-24372.75
35	-10285.25	-8923.18	84.26	1173.68	-2219.48	-20169.97
36	-22319.71	-18089.79	-2479.32	-1848.59	-1658.37	-46395.78
38	-40122	-4609.37	623.25	-19603.4	34786.49	-28925.03
40	3602.46	-6215.82	-20250.59	835.61	768.47	-21259.87
41	-6163.57	-8547.37	-21957.44	213.62	-315.93	-36770.69

To sum up, now that China has made clear the goal of “double carbon”, how to quickly unlock the carbon bondage has become the key to the next step. Future scholars can apply this method to the analysis of other regions or other industries to accurately eliminate the high dependence on fossil energy. Based on this, the government should refine

the specific targets of carbon emissions by industry segments and build a green energy economic system, so as to reduce the consumption of fossil fuels. All industries should take the initiative to implement energy conservation and emission reduction and technological upgradings, such as carbon capture, storage, or negative emission technologies, so as to

prevent carbon emissions from rising again and realize the two parallel rounds of optimizing the industrial structure and realizing the goal of “double carbon”.

### Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

### Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

### Acknowledgments

This work was financially supported by the Research on Promotion Strategy of Customer Resources to Competitiveness Service Enterprises in Liaoning (Social Science Planning Fund of Liaoning Province, L19BGL029).

### References

- [1] R. Pincheira, F. Zúñiga, and F. Valencia, “An environmental measurement for a dynamic and endogenous global environmental Kuznets curve in the global context,” *Environmental Science and Pollution Research*, vol. 28, no. 46, Article ID 65594, 2021.
- [2] Z. W. Yu, Y. P. Fan, and H. Luo, “Research on the impact of China's advanced industrial structure on carbon emission intensity,” *East China Economic Management*, vol. 36, no. 1, pp. 78–87, 2022.
- [3] A. Mujtaba, P. K. Jena, and D. Mukhopadhyay, “Determinants of CO<sub>2</sub> emissions in upper middle-income group countries: an empirical investigation,” *Environmental Science and Pollution Research International*, vol. 27, no. 30, Article ID 37759, 2020.
- [4] S. Georgia, “Royal Court sets out plan to become carbon neutral,” *Stage*, vol. 11, p. 5, 2020.
- [5] F. J. André and J. A. Valenciano-Salazar, “Becoming carbon neutral in Costa Rica to Be more sustainable: an AHP approach,” *Sustainability*, vol. 12, no. 2, p. 737, 2020.
- [6] E. Demetriou, G. Mallouppas, and C. Hadjistassou, “Embracing carbon neutral electricity and transportation sectors in Cyprus,” *Energy*, vol. 229, 2021.
- [7] Anonymous, “California cement producers tie carbon neutral operation to sound policies,” *Concrete Products*, vol. 124, no. 5, pp. 16–17, 2021.
- [8] R. K. Srivastava, N. P. Shetti, K. R. Reddy, E. E. Kwon, M. N. Nadagouda, and T. M. Aminabhavi, “Biomass utilization and production of biofuels from carbon neutral materials,” *Environmental Pollution*, vol. 276, Article ID 116731, 2021.
- [9] M. L. Jeffery, J. Gütschow, R. Gieseke, and R. Gebel, “PRI-MAP-crf: UNFCCC CRF data in IPCC 2006 categories,” *Earth System Science Data*, vol. 10, no. 3, pp. 1427–1438, 2018.
- [10] National Technical Committee for standardization of energy infrastructure and management (SAC/TC 20), “General Rules for Calculation of Comprehensive Energy Consumption,” 2020.
- [11] B. Navdeep, Y. Selin, S. Z. Jibrán, E. Wolfgang, and K. P. Martin, “The evolution of energy efficiency in

Switzerland in the period 2000–2016,” *Energy*, vol. 191, C, Article ID 116526, 2020.

- [12] H. C. Chin, X. T. Wei, J. T. Zhao et al., “The driving factors of energy-related CO<sub>2</sub> emission growth in Malaysia: the LMDI decomposition method based on energy allocation analysis,” *Renewable and Sustainable Energy Reviews*, vol. 115, Article ID 109356, 2019.

## Research Article

# A Risk Decision-Making Model Based on Kalman Filter for Intellectual Property Pledge Financing

Xianan Yin , Hua Ming , and Xinzhong Bao 

Management College, Beijing Union University, Beijing 100101, China

Correspondence should be addressed to Xinzhong Bao; [baoxz@buu.edu.cn](mailto:baoxz@buu.edu.cn)

Received 17 December 2021; Accepted 20 January 2022; Published 27 February 2022

Academic Editor: Daqing Gong

Copyright © 2022 Xianan Yin et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Risk dynamic early warning is of great importance for financing risk decision-making. Intellectual property (IP) pledge financing is an effective way to alleviate the financial difficulties for technologically small- and medium-sized enterprises (SMEs). It is very important to study the financing risk decision-making because of its higher risk compared with other mortgage loans. Based on Kalman filter, we establish the pledge financing risk decision-making model and extract the key variables affecting financing risk by principal component analysis. We test the model with 88 listed SMEs. The results show that the average error between the predicted and the real values is 8.5% and the overall recognition accuracy of the model is 89.1%. The risk decision-making model has high discriminant accuracy and can provide evidence to risk decision-making.

## 1. Introduction

Technology SMEs are counted as one of the most technologically advanced enterprises that assign great large amount of money to their development [1, 2]. Furthermore, intellectual property pledge financing is an effective way to alleviate the financing difficulties of technology SMEs. As the significant differences between IP pledges and traditional physical pledges lead to specificity in both asset valuation risk and disposal risk, risk is a primary consideration when stakeholders make financing decisions. However, the risk is an uncertain event or condition that, if it occurs, has a negative effect, and it could happen at any stage without warning [3]. Therefore, many scholars have studied the IP pledge financing risks in different aspects. Pennington & Sanchez (2007) and Crawford & Strasser (2008) studied the IP pledge financing risks in different stages from three dimensions: credit risk, infringement risk, and compensation risk [4, 5]. Moreover, from the essence of IP pledge financing risk, in addition to the quality of the pledged goods, the stakeholders of pledge financing pay more attention to the financial and operational conditions of the financing enterprises. Liu Peipei (2011) believes that the risk of IP pledge financing depends on the value of the pledged goods

formally, but on the overall business capacity of enterprises essentially [6], which has also been shared by most scholars. While the financial status and operating conditions of an enterprise are a dynamic process that develops and changes in real time, a dynamic risk decision model could better characterize the risk of IP pledge financing.

The current research on the risk of IP pledge financing mainly focuses on the construction of a financing risk index system [7–9], while the research on financing risk decisions is limited to static evaluation and analysis. In addition, an advanced tool for validating decisions is necessary to ensure the smooth operation of corporate activities [10]. Although common methods can be used to evaluate financing risks, including factor analysis method [11], VIKOR method [12], interval number TOPSIS [13], hierarchical analysis method, and fuzzy comprehensive evaluation method [14], they mainly use the cross-sectional data, which cannot reflect the cumulative formation process of risks, so the models of traditional methods are not stable enough, which reduces the accuracy of risk decision-making. Besides, decision-making action that effectively extracts and utilizes enough data about decision based on big data is a trend [15].

In order to improve the effectiveness of risk decision-making, scholars have tried to introduce a dynamic risk



decision model called BP neural network that calculates the error between the calculated output value and the true value, and continuously adjust it until the acceptable error level [16, 17]. Zeng and Ming evaluate the risk of IP pledged financing for technology-based SMEs empirically [18]. BP neural network achieves early warning of dynamic risk with the deep learning algorithms, but it is essentially an approximation algorithm for linear weight functions, which is more effective if there is no correlation among processing data. As for the strong correlation of time series data of business operations, it would reduce the effectiveness of early warning of dynamic risk. Therefore, IP pledge financing risk requires methods that can analyze data with correlated time series and reflect the risk formation process, and Kalman filtering meets these needs well. Compared with other methods, the Kalman filter has two advantages. First, the correlation among processing data will not affect the results. Second, the filtering process is an acyclic recursive process of automatically updating “prediction-correction” with the latest data, and it does not need to save the historical measurement data. The enterprise’s IP pledge financing risk is a dynamic state with enterprise operation, and the previous risk will affect the status of the next period state. Therefore, the Kalman filtering method can predict the dynamic financing risk and determine the level of financing risk.

Kalman filtering has been widely used in the field of management for monitoring risk [19], especially for dynamic risk. Sun et al. established a Kalman filter dynamic early-warning model for financial crises, and the results showed that Kalman filter-based dynamic model is more accurate than other prediction models [20]. Zhu studied the financial crises in different life cycle stages of enterprises based on the dynamic model of Kalman filter combined with the logistic regression method [21]. Fu et al. constructed an early-warning model of financial risk with the fusion of neural network and Kalman filter and verified that enterprise financial index data has a good predictive effect on the enterprise financial risk [22].

In summary, the Kalman filter theory is more applied to the dynamic early warning of enterprise financial risk, not to the financing risk decision-making. Since the Kalman filter can describe the cumulative process of risk change and provide risk early warning by the time series data, we will try to establish a financing risk decision-making model based on Kalman filter and test the accuracy and effectiveness of the model.

The possible contributions of this paper are mainly reflected in two aspects: (1) We construct a dynamic decision-making model instead of the previous static model, which improves the accuracy of risk decision. (2) We introduce a risk decision-making method which is not affected by time series correlation data. The Kalman filtering method could eliminate the adverse effects of the correlation data. According to the dynamics and accumulation of financing risk, we use the self-cyclic recursive characteristics of Kalman filter to identify the parameters of the model and estimate the risk status with the time series financial data, and visualize the financing risk level.

## 2. A Risk Dynamic Decision-Making Model

*2.1. Kalman Filtering.* Kalman filter is a method based on state-space model, which can estimate the hidden state of the system in real time through the observation with noise, so as to solve the problem of state-space estimation. The method can also include any number of unknowns; it treats the signal process as the output of a linear system under the action of white noise and uses the state equation to describe this input-output relationship, using the statistical properties of the system state equation, the observation equation, and the white noise excitation (system noise and observation noise) to form a filter with recursive characteristics algorithm. The core consists of the following five equations:

State updating equation: the state estimation of the current time point from the previous state is as follows:

$$\hat{x}_t = A_t \hat{x}_{t-1} + u_t. \quad (1)$$

The calculation of the estimated covariance of the current time point state is as follows:

$$P_t^- = A_t P_{t-1} A_t^T + Q_t. \quad (2)$$

Measuring updating equation: the calculation of Kalman gain is as follows:

$$K_t = \frac{P_t^- H_t^T}{H_t P_t^- H_t^T + R_t}. \quad (3)$$

$y_t$  is measured to update the state estimation.

$$\hat{x}_t = \hat{x}_t^- + K(y_t - H_t \hat{x}_t^-). \quad (4)$$

Updating error covariance is as follows:

$$P_t = (I - K_t H_t) P_t^-. \quad (5)$$

The state update equation estimates the state of the current time point using the estimated value obtained from the previous time point and the covariance of the error to obtain a priori estimate. On the other hand, the measurement update equation combines the obtained a priori estimate with the measured value to obtain an improved a posteriori estimate, which serves as feedback to the state update equation and continues to estimate the next time point. Thus, a recursive process of “prediction-correction” is formed. Among them, Kalman gain plays a role in adjusting the weight between the observed value and the predicted value, so as to make the posterior estimate closer to the real value.

*2.2. Establishing State-Space Model for Risk Decision-Making.*

In this equation,  $X_t$  is composed of random variables  $x_t$ , representing the financing risk status of financing enterprises in the period, and  $Y_t$  is composed of observation values of various specific indicators  $y_t$ , representing the principal component data after indicator-specific data processing. Assuming that the operating conditions of financing enterprises in each year are regarded as a discrete

control process system and their real operating conditions cannot be observed, but can be predicted by the relationship between the observed values and white noise incentives, the state-space model of dynamic early warning of financing risks can be expressed by the following state equation and observation equation, respectively:

$$x_t = A_t x_{t-1} + B_t u, \quad (6)$$

$$y_t = H_t x_t + v. \quad (7)$$

The parameter vector  $A_t$  is estimated according to the historical data of the operation status of the financing enterprises. The parameter  $H$  used to measure the system is the value of the measurement index of the financing risk of the enterprises and corresponds directly to the output results.  $u_t$  and  $v_t$  represent the process noise and the measurement noise of the system, respectively, being a white noise sequence:

$$u_t \sim N(0, Q_t), \quad (8)$$

$$v_t \sim N(0, R_t). \quad (9)$$

The spatial state model of pledge financing risk is a single model and single measurement, and with the advance of filtering, it has convergence, so the initial value is set to 0. This dynamic early-warning model of IP pledge financing risk based on Kalman filter is constructed by time series data as the following equations.

$$Y_t = (11)X_t, \quad (10)$$

$$X_t = \begin{pmatrix} 1 & 0 \\ 0 & A \end{pmatrix} X_{t-1} + \begin{pmatrix} u_t \\ v_t \end{pmatrix}, \begin{pmatrix} u_t \\ v_t \end{pmatrix} \sim N(0, H). \quad (11)$$

$A$  and  $H$  will be updated continuously with the promotion of the financial data filtering of the test samples in the equations.

### 3. Methodology

**3.1. Data.** According to the characteristics of the IP pledge financing enterprises that have been obtained from the official website of the State Intellectual Property Office in China, it can be seen that most of the IP pledge financing enterprises are small- and medium-sized technological enterprises. Accordingly, we choose the listed companies of technological SMEs as samples and exclude the following: (1) ST and \*ST enterprises, which have unrepresentative debt repayment ability and abnormal financial data; (2) financial insurance enterprises; (3) enterprises with incomplete data disclosure. Combining the requirements of Kalman filtering and the characteristics of loan repayment risk, this paper uses net profit and net operating cash flow as two indicators to measure the degree of financing risk of enterprises, and 88 sample enterprises are selected. The less the interval data of Kalman filtering model, the better the filtering effect, so quarterly and semiannual reports should be the appropriate data for empirical analysis. However,

because of the lack of quarterly data, we use the financial data of semiannual reports from 2013 to 2016. Data mainly comes from the WIND database, and individual index data is obtained by searching the financial reports of listed companies.

**3.2. Constructing Early-Warning Index System of Dynamic Risk.** The IP pledge financing risk is mainly reflected in the financial risk and operational risk of enterprises. The influence of nonfinancial information in the operation status of enterprises will be concentrated on the financial indicators of enterprises which contain a lot of information and can be used to predict the operation status of enterprises [23]. There are obvious differences in the financial indicators between the enterprises with financial crisis and enterprises with health financial position [24]. Before the company's financial crisis begins to worsen, the relevant characteristics will be reflected in financial indicators in advance [25]. Whether the financial indicators are calculated directly or indirectly through financial reports, they can be used effectively to study the financial risk [26]. Therefore, following the principles of scientificity, comparability, and data availability, we added the indicators of cash flow and innovation ability on the basis of traditional financial risk measurement dimensions of solvency, profitability, operation ability, and growth ability to test the risk of IP pledge financing. As shown in Table 1, a risk early-warning evaluation index system of pledge financing with 21 indicators in 6 dimensions is constructed.

**3.3. Global Principal Component Analysis of Dynamic Data.** We obtained 616 sets of multidimensional time series stereo data from 88 families in 7 periods. In order to quickly extract important information from the stereo data table and analyze the dynamic law of the system, the global principal component analysis method was used to reduce the dimension of indicators, and several interrelated numerical indicators were transformed into a few independent comprehensive indicators; that is, fewer indicators are used instead of multiple indicators to combine most of the original information. We used the global principal component method to analyze the three-dimensional data table by SPSS 20.0. The total variance and coefficient matrix are shown in Tables 2 and 3.

In order to obtain the most orthogonal principal component factors while retaining most of the information and to get the optimal input database, the first 15 global principal component factors were extracted. The cumulative contribution rate reached more than 94%, which means the effect was good. These global principal component factors are all represented by a linear combination of 21 financial indicators. According to the coefficient matrix, the formula of each principal component is shown as

$$F_n = \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \dots + \alpha_{20} x_{20} + \alpha_{21} x_{21}. \quad (12)$$

Among them,  $n$  is the number of extracted principal components ( $n = 1, 2, 3 \dots 15$ ),  $\alpha$  is the coefficients of each

TABLE 1: Early-warning index system.

Type	Code	Indicator name
Debt paying ability	$X_1$	Current ratio
	$X_2$	Quick ratio
	$X_3$	Asset-liability ratio
	$X_4$	Cash ratio
	$X_5$	Cash-maturity debt ratio
Cash flow	$X_6$	Operating income to sales cash ratio
	$X_7$	Operating income to net cash ratio
	$X_8$	Cash operation index
Financial performance	$X_9$	Return on equity
	$X_{10}$	Profit on asset
	$X_{11}$	Net profit on asset
	$X_{12}$	Sales of net profit margin
Innovation capability	$X_{13}$	Retained earnings on asset
	$X_{14}$	R&D-revenue ratio
	$X_{15}$	Number of bachelor degrees or above in staff
Asset operation	$X_{16}$	Inventory turnover
	$X_{17}$	Receivable turnover
	$X_{18}$	Asset turnover
Growth ability	$X_{19}$	Net profit growth ratio
	$X_{20}$	Revenue growth ratio
	$X_{21}$	Net asset growth ratio

TABLE 2: Total variance explained.

Component	Initial eigenvalue			Extraction sums of squared loadings		
	Total	% of variance	Cumulative %	Total	% of variance	Cumulative %
1	4.221	20.102	20.102	4.221	20.102	20.102
2	2.996	14.268	34.370	2.996	14.268	34.370
3	1.693	8.063	42.433	1.693	8.063	42.433
4	1.340	6.380	48.813	1.340	6.380	48.813
5	1.304	6.208	55.021	1.304	6.208	55.021
6	1.162	5.534	60.555	1.162	5.534	60.555
7	1.045	4.976	65.532	1.045	4.976	65.532
8	0.998	4.753	70.285	0.998	4.753	70.285
9	0.891	4.244	74.529	0.891	4.244	74.529
10	0.848	4.039	78.568	0.848	4.039	78.568
11	0.819	3.902	82.470	0.819	3.902	82.470
12	0.754	3.593	86.062	0.754	3.593	86.062
13	0.708	3.372	89.434	0.708	3.372	89.434
14	0.572	2.722	92.156	0.572	2.722	92.156
15	0.477	2.274	94.430	0.477	2.274	94.430
16	0.442	2.105	96.535			
17	0.348	1.655	98.190			
18	0.302	1.436	99.626			
19	0.063	0.299	99.925			
20	0.008	0.040	99.965			
21	0.007	0.035	100.000			

evaluation index, and  $x$  refers to the standardized values of each evaluation index.

For example, the values for each of the sample companies in period  $F_1$  can be expressed as

$$\begin{aligned}
F_1 = & 0.53x_1 + 0.546x_2 - 0.676x_3 + 0.515x_4 + 0.186x_5 - 0.032x_6 \\
& + 0.339x_7 + 0.078x_8 + 0.7x_9 + 0.808x_{10} + 0.844x_{11} + 0.506x_{12} \\
& + 0.425x_{13} + 0.196x_{14} + 0.215x_{15} - 0.142x_{16} - 0.002x_{17} + 0.084x_{18} \\
& + 0.29x_{19} + 0.43x_{20} + 0.301x_{21}.
\end{aligned} \tag{13}$$

TABLE 3: Coefficient matrix.

	Component														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$X_1$	0.530	-0.798	0.193	0.012	-0.094	-0.051	-0.067	0.059	-0.035	-0.025	-0.007	-0.054	-0.027	-0.011	-0.033
$X_2$	0.546	-0.791	0.206	-0.014	-0.060	-0.035	-0.045	0.087	-0.048	-0.015	0.022	-0.057	-0.009	0.015	-0.003
$X_3$	-0.676	0.198	0.216	0.166	0.131	-0.137	-0.075	0.040	-0.010	0.047	0.260	-0.245	0.155	0.070	0.139
$X_4$	0.515	-0.755	0.251	-0.013	-0.054	-0.040	-0.035	0.131	-0.040	0.026	0.047	-0.060	0.012	0.024	0.038
$X_5$	0.186	0.086	-0.127	-0.147	0.481	0.188	0.089	0.529	-0.301	-0.011	0.482	0.047	0.134	-0.043	-0.085
$X_6$	-0.032	0.022	0.271	0.164	0.497	-0.391	0.419	0.200	0.102	0.049	-0.428	-0.037	0.160	-0.123	-0.169
$X_7$	0.339	0.141	0.227	-0.563	0.065	0.274	0.192	-0.040	0.064	0.349	-0.192	-0.093	0.314	0.114	0.267
$X_8$	0.078	0.063	0.303	-0.410	0.037	0.037	0.434	-0.128	0.378	-0.539	0.249	-0.048	-0.154	0.009	0.016
$X_9$	0.700	0.381	-0.033	0.015	0.072	-0.155	-0.146	-0.223	-0.050	-0.016	0.072	-0.052	0.001	-0.048	0.048
$X_{10}$	0.808	0.446	0.043	-0.021	0.025	-0.078	-0.078	-0.116	-0.100	-0.029	0.034	-0.026	0.056	-0.212	0.055
$X_{11}$	0.844	0.397	-0.003	-0.017	0.023	-0.081	-0.071	-0.112	-0.089	-0.030	0.014	0.021	0.026	-0.201	0.065
$X_{12}$	0.506	0.215	0.035	0.096	0.103	-0.571	-0.058	-0.078	0.045	-0.029	0.140	0.146	0.096	0.491	-0.063
$X_{13}$	0.425	0.227	-0.273	-0.121	0.337	0.238	-0.045	0.221	-0.068	-0.047	-0.313	-0.029	-0.495	0.240	0.005
$X_{14}$	0.196	-0.342	-0.258	0.142	0.305	0.344	0.109	-0.356	0.243	0.160	0.112	0.482	0.156	0.034	-0.123
$X_{15}$	0.215	-0.217	-0.125	0.612	0.430	0.154	0.108	-0.210	0.134	-0.031	0.075	-0.318	-0.090	-0.089	0.214
$X_{16}$	-0.142	0.124	0.596	0.193	0.230	0.181	-0.435	0.194	0.208	-0.122	-0.068	0.245	0.009	0.108	0.257
$X_{17}$	-0.002	0.138	0.433	0.295	-0.176	0.153	0.505	-0.192	-0.457	0.177	0.096	0.064	-0.192	0.193	0.066
$X_{18}$	0.084	0.359	0.728	0.102	-0.006	0.241	-0.161	-0.053	-0.051	0.008	-0.018	0.088	-0.072	-0.148	-0.316
$X_{19}$	0.290	0.243	0.032	0.119	-0.260	-0.131	0.130	0.330	0.488	0.500	0.238	0.030	-0.276	-0.092	0.012
$X_{20}$	0.430	0.276	-0.069	0.257	-0.246	0.419	-0.025	0.099	0.200	-0.075	-0.035	-0.390	0.269	0.243	-0.258
$X_{21}$	0.301	0.177	-0.180	0.379	-0.408	0.069	0.269	0.364	-0.027	-0.311	-0.140	0.294	0.191	-0.050	0.192

The value of  $F$  means that the financing risk was represented by 15 principal components instead of 21 original evaluation indicators, and the interpretation rate can reach 94%. Due to the space constraints, we only listed the formula of  $F_1$  as sample. The result data will be used for the next filtering calculation.

## 4. Results

**4.1. Dividing Risk Level.** Although the risk of IP pledge financing is an asymptotic dynamic process of the enterprise's financial situation from good to bad, the degree of risk of IP pledge financing is different at the early-warning level. Generally speaking, the most serious degree of financial risk is bankruptcy due to insolvency, while the most serious degree of IP pledge financing risk is the lack of enough cash to repay the financing money due, but far from liquidation bankruptcy. Therefore, besides the net profit, the most important factor in measuring the risk of pledge financing is adequacy of cash flow. Consequently, in this paper, we defined the degree of risk of IP pledge financing from two indicators: net profit and net operating cash flow. The specific level of risk is divided according to the following standards:

**Healthy enterprises:** If the net profit and net operating cash flow of enterprises are both positive during the investigation period and show an increasing trend, the enterprises are identified as healthy ones. There are 44 healthy enterprises in the 88 samples, of which 28 are forecast samples and 16 are test samples.

**Mild-risk enterprises:** If the net cash flow of enterprises is negative at the end of the inspection period, the net cash flow is positive at T-1 and T-2 periods, and the net profit of any period is negative during the inspection period, the enterprises are identified as mild-risk ones. There are 24 mild-risk

enterprises in the 88 samples, of which 16 are forecast samples and 8 are test samples.

**High-risk enterprises:** According to the data of listed companies on SME technology boards, only 11 of the 75 samples with negative net operating cash flow in three consecutive periods have positive net operating cash flow in 2015. That is, if an enterprise has negative net cash flows in periods T and T-1, it has an 85% probability of negative net cash flows in period T-2, which shows that the net operating cash flow has certain sustainability. Therefore, we identified the degree of high risk which is the negative value of net operating cash flow and twice the negative value of net profit in three consecutive periods. There are high-risk 20 enterprises among the 88 samples, of which 12 are forecast samples and 8 are test samples.

In summary, the 88 sample companies are divided into two groups: the first group is the forecast sample group, which consists of 56 companies, including 28 healthy companies and 28 risky companies, and the second group is the test sample group, which consists of 32 companies, including 16 healthy companies and 16 risky companies. Predictive samples are used as training sets to optimize the model, and test samples as detection sets to verify the effectiveness of risk decision-making model.

**4.2. Defining Risk-Level Threshold.** In this study, statistical analysis was used to extract judgment thresholds for the risky companies based on the data from the forecast sample. With a 95% confidence probability and a confidence coefficient of 1.6449, a mean value of 0.056 and a standard deviation of 0.016 were calculated for the risky sample. The lower confidence limit for a crisis in the risky sample of companies is as follows:



$$\text{down} = \text{mean} - \alpha \times \text{Standard Deviation} = 0.056 - 1.6449 \times 0.016 = 0.030. \quad (14)$$

Similarly, under the precondition of 95% confidence probability, the confidence coefficient is 1.6449, the mean value of the calculated state for healthy samples is 0.063, and

$$\text{up} = \text{mean} + \alpha \times \text{Standard Deviation} = 0.063 + 1.6449 \times 0.007 = 0.075 \quad (15)$$

According to the above results, the lower confidence limit is 0.030, and the upper confidence limit is 0.075. That is, when the predicted value of financing risk is less than 0.030, the high financing risk probably happens; when the predicted value of financing risk is greater than 0.075, the financing enterprises are in a healthy state; and when the predicted value of financing risk is between 0.030 and 0.075, the financing enterprises are likely to be in a state of mild financing risk.

**4.3. Testing the Dynamic Risk Decision-Making Result.** According to the global principal component analysis, we have obtained 15 principal components' factor loads and constructed a comprehensive index of the company's financial situation in each period from the characteristic value  $F$  and contribution rate  $\beta$  (Table 2), which is the linear combination of each principal component. We have obtained a general index  $X_t$  reflecting the financing risk status of each half year of the test sample, and the results are shown in Table 4.

$$Z = \beta_1 F_1 + \beta_2 F_2 + \beta_3 F_3 + \cdots + \beta_{14} F_{14} + \beta_{15} F_{15}. \quad (16)$$

Inputting the data of 32 test samples from 7 periods into the Kalman filtering model, we obtained the predicted values of each year. The results of some samples are shown in Table 4.

Due to the space constraints, we choose four enterprises' dynamic forecasting charts of financing risk, including two healthy enterprises and two risky enterprises, in Figure 1.

**4.4. Result Analysis.** The results of Table 4 and Figure 1 show that the dynamic pledge financing risk decision-making model based on Kalman filter is a recursive updating process. After the initial estimation of the formation state, the initial estimation is modified by inputting the updated data information every year, and the stable state of the predicted value and the real value is gradually achieved. The model is based on the cumulative deviation of the financial situation of an enterprise over time to represent the evolution process of the financing risk of an enterprise. In a certain period of time, the financial situation of a healthy enterprise is basically in a good category, and the trend of change is gradually better, but it does not exclude a temporary weakening state. The financing risk of a risky company is gradually getting worse from mild crisis to severe crisis. There are

the standard deviation is 0.007. The upper confidence limit for healthy sample companies in crisis is as follows:

significant dynamic early-warning characteristics to predict the financing risk by cumulative variation rather than a cross-sectional state. In addition, we can determine whether the financing risk is temporary crisis or continuous deterioration in the shape of the forecast chart and observe the change point of the financing risk state, so as to realize the visual effect of financing risk decision-making. Next, we choose one of the healthy samples and risky samples to illustrate the forecast chart.

We take Yunnan Luoping Zinc & Electricity Co., Ltd. (stock code: 002114) as an example of a healthy enterprise. The enterprise was founded in 2000 and is a high-tech enterprise. Its main business is lead-zinc ore mining, zinc smelting, hydroelectric power generation, and comprehensive utilization of resources. It also has an integrated industrial chain of "mine-electricity-smelting." At present, it has successfully developed a number of advanced production technologies for comprehensive utilization of resources, such as germanium-indium extraction, purification of workshop slag treatment, leaching slag flotation of silver, and zinc powder mill pulverizing process transformation. From the forecast chart, only the sixth observation period is slightly lower than the healthy threshold, while other observation periods are above the healthy threshold, indicating that the company's overall solvency is strong. The annual value in 2015 is much lower than that in 2014 for the reason that in the second half of 2015, due to the steady recovery of the US economy and the expected increase in interest rates, the high domestic zinc price has fallen back and continued to fall. Although it adapted to the rapid adjustment of its business ideas in the market, it is still affected by the overall business risk and financial risk. In 2016, due to the influence of such favorable factors as "capacity removal, inventory removal, deleveraging, cost reduction, and shortage compensation board," the decline of zinc price eased and slowly rebounded. It can be seen from the forecast chart that although the actual value is low, the forecast value is above the healthy threshold line, and the actual value at the end of 2016 is much higher than the healthy threshold line, which shows the accuracy of the model. The actual value and forecast value of other prediction points are in the same healthy area, and the change direction of forecast value is consistent with the change direction of actual value in the next period, which further proves that Kalman filter model has good tracking and early warning of the risk of pledge financing.

We take Carrier Holdings Co., Ltd. (stock code: 002072) as an example of risk-based enterprises. The enterprise is a cotton

TABLE 4: The real values and the predictive values of sample companies.

Data	Real value	Predictive value	Real value	Predictive value	Real value	Predictive value
Company code		002114		002098		002450
2013.12.31	0.1716	0.1716	0.0691	0.0691	0.1049	0.1049
2014.06.30	0.1013	0.1540	0.0365	0.0609	0.0616	0.0940
2014.12.31	0.2328	0.1994	0.0653	0.0600	0.0930	0.0882
2015.06.30	0.2366	0.2348	0.0303	0.0369	0.0529	0.0597
2015.12.31	0.1252	0.1608	0.0684	0.0565	0.1014	0.0857
2016.06.30	0.0667	0.0810	0.0413	0.0467	0.0619	0.0700
2016.12.31	0.1647	0.1286	0.0856	0.0752	0.0932	0.0858
Company code		002139		002020		002014
2013.12.31	0.1037	0.1037	0.0930	0.0930	0.0936	0.0936
2014.06.30	0.0517	0.0907	0.0657	0.0862	0.0415	0.0806
2014.12.31	0.1030	0.0929	0.1001	0.0924	0.0914	0.0818
2015.06.30	0.0554	0.0641	0.0531	0.0636	0.0467	0.0546
2015.12.31	0.1097	0.0935	0.0834	0.0740	0.1006	0.0845
2016.06.30	0.0535	0.0663	0.0523	0.0578	0.0543	0.0648
2016.12.31	0.1096	0.0957	0.0839	0.0750	0.0978	0.0877
Company code		002072		002125		002260
2013.12.31	-0.0388	-0.0388	0.0612	0.0612	0.1104	0.1104
2014.06.30	-0.1847	-0.0753	0.0112	0.0487	0.0094	0.0852
2014.12.31	0.0553	-0.0036	-0.0421	-0.0192	0.0933	0.0787
2015.06.30	-0.0402	-0.0198	0.0168	-0.0066	0.0299	0.0388
2015.12.31	-0.2683	-0.2051	0.0681	0.0466	0.0369	0.0298
2016.06.30	0.0675	-0.0284	0.0552	0.0605	0.0318	0.0272
2016.12.31	0.0032	0.0136	0.0491	0.0567	-0.0286	-0.0168
Company code		002392		002571		002115
2013.12.31	0.0578	0.0578	0.0607	0.0607	-0.0064	-0.0064
2014.06.30	0.0276	0.0502	0.0214	0.0509	-0.1934	-0.0531
2014.12.31	0.0399	0.0396	0.0386	0.0378	0.0500	-0.0051
2015.06.30	0.0131	0.0170	0.0122	0.0153	0.0375	0.0325
2015.12.31	0.0071	0.0050	-0.0341	-0.0266	0.0700	0.0688
2016.06.30	0.0125	0.0071	-0.0395	-0.0447	-0.0429	-0.0069
2016.12.31	0.0515	0.0398	0.0913	0.0524	0.0489	0.0262

textile enterprise listed in 2006. Its main products are extra wide decorative fabric series, large and small jacquard fabric series, and so on. It is at high risk according to the risk classification criteria for pledge financing of the original samples. From the results of the forecast chart of the observation period, the enterprise is in the area of lower risk threshold for most of the time, and the overall situation is a high risk.

In 2014, affected by many factors, such as the downturn of domestic real economy and the rising price of factors of production, cotton textile enterprises had increased operating pressure causing domestic demand to slump continuously and its international market competitiveness to be declined significantly; in addition, the development of the textile industry has become more difficult. From individual observation points, the true values of the third and sixth phases are above the lower limit of the risk threshold, but the predicted values are in the severe risk area, and the subsequent fourth and fifth phases return to the severe risk area, thus confirming the accuracy of the decision-making results.

*4.5. Total Discrimination Accuracy Analysis.* The overall discriminatory accuracy analysis is shown in the following three areas:

- (1) Overall prediction accuracy analysis. Any risk decision-making model has its limitations or neglected influencing factors. Therefore, there is no perfect model which achieves 100% prediction accuracy. Prediction accuracy refers to the goodness or disadvantage of the fitting degree between the predicted value and the actual value produced by the prediction model, which reflects the fitting degree of the prediction model, and is also an important index to judge whether the prediction method has applicability. According to the test results, the average error between the predicted value and the real value of the seven-period semiannual reports is 8.5%, and the average error of the model in the last period is only 5.1%. Accordingly, the prediction accuracy of the model is high and the prediction effect is good.
- (2) Analysis of the results of risk-based test samples. From the predicted value calculated by the dynamic model of Kalman filter, if the last stage of the whole sample is risk-based, and 93% of the test samples in the sixth period are correctly predicted before the crisis occurs, that is, before the risk situation appears at the end of 2016, they have already fallen below the threshold of mild crisis or entered the area of high risk, and 93% of



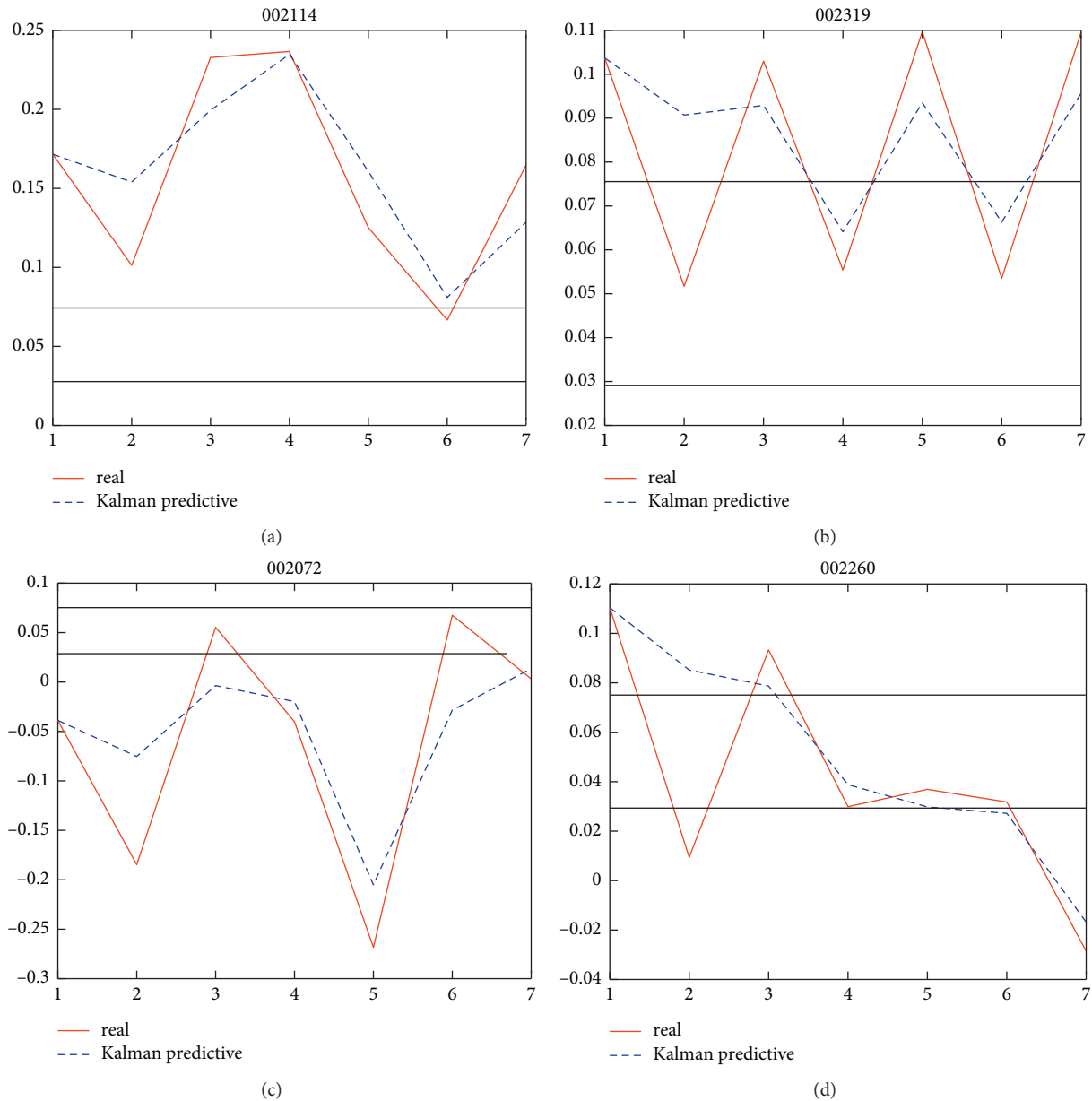


FIGURE 1: Some of the predictive value curves of testing the financing risk.

the samples have the mild financing risk two times in advance.

- (3) Analysis of the results of healthy-type test samples. From the predicted results of healthy samples, the predicted values of most healthy enterprises are above the warning line of mild-risk threshold. The predicted values of individual samples have temporarily deviated from the normal values at a certain point, but the dynamic model then corrects the situation in time. Even if they belong to mild-risk area, the predicted values of some healthy enterprises also deviate from the normal values, which means they are slightly below the warning line and are at a temporary mild risk. The results of 112 groups of dynamic

data from 16 healthy enterprises showed that the recognition accuracy of healthy samples in the sixth period was 93.75% and the overall recognition accuracy was 87.5%.

This paper further classifies the errors in classification recognition: class A and class B errors. When a risky financing enterprise is misjudged as a healthy enterprise, this is called class A error. If it is correctly identified, it is considered to be sensitive. Conversely, when a healthy enterprise is misjudged as a risky enterprise, this is called class B error. If it is correctly identified, it is considered to be specific. No matter what kind of mistake occurs, it may lead to wrong decision-making and cause serious losses. The accuracy of decision-making model in identifying risk categories is shown in Table 5.

TABLE 5: Recognition accuracy of Kalman filter model.

Date	Class A error ratio (%)	Sensitivity (%)	Class B error ratio (%)	Specificity (%)	Recognition accuracy (%)
2016.12	6.25	93.75	6.25	93.75	93.75
2016.06	12.5	87.5	6.25	93.75	90.6
2015.12	6.25	93.75	12.5	87.5	90.6
2015.06	6.25	93.75	18.75	81.25	87.5
2014.12	12.5	87.5	12.5	87.5	87.5
2014-06	12.5	87.5	18.75	81.25	84.3

According to the results in Table 5, the Kalman filter model has a high comprehensive judgment rate, which shows that the Kalman filter model has good robustness and prediction ability. The rate of misjudging a risky company as a healthy enterprise is 9.4%, the rate of misjudging a healthy company as a risky company is 12.5%, and the overall accuracy of the model is 89.1%. Among them, the rate of risky enterprises misjudged as healthy ones is lower than that of healthy enterprises misjudged as risky ones, which can further reduce the risk of pledge financing.

In addition, the results also illustrate that the risk criteria for IP pledge financing are likely to be more stringent than those of other financing ways. For banks and other lending institutions, it is necessary to change the previous cautious attitude towards IP pledge financing business.

## 5. Conclusion

Dynamic prediction can provide sufficient evidence for financing risk decision-making and reduce financing decision risk. In this paper, we established the pledge financing risk decision-making model based on Kalman filter, because of cumulative and time-varying characteristics of financing risks, and extracted the key variables affecting financing risk by principal component analysis method. Listed SMEs are tested and analyzed, and we found the following: (1) The average error between the predicted and the real values is 8.5%, so the model has higher fitting precision. (2) The rate of a risky company wrongly judged to be a healthy company is 9.4%. Conversely, the rate is 12.5% for a healthy company wrongly judged to be a risky company. The overall recognition accuracy of the model is 89.1%. (3) The risk decision-making model has high discriminant accuracy and can provide evidence to risk decision-making.

In this paper, we only established the pledge risk decision-making model based on Kalman filter. As we all know, the risk of IP pledge is a very complex problem. In future research, we will explore the establishment of a risk decision-making model based on combination of Kalman filter and neural network, so as to further improve the accuracy and effectiveness of the risk decision-making model.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This research was supported by Major Projects of Beijing Social Science Foundation (no. 20ZDA03).

## References

- [1] M. Ramezani and M. Khazaei, "Internationalization of small and medium enterprises (SME's): dimensions and strategies," *Journal of Logistics, Informatics and Service Science*, vol. 7, no. 2, pp. 65–96, 2020.
- [2] I. P. Doina, C. Sebastian-Ion, A. Adriana, and C. Eduard-Gabriel, "Relationships between knowledge absorptive capacity, innovation performance and information technology," *Case study: The Romanian Creative Industries SMEs, Studies in Informatics and Control*, vol. 28, no. 4, pp. 463–476, 2019.
- [3] M. A. Safayet, M. H. Islam, and S. Ahmed, "A case study on risk management in existing construction project in Bangladesh," *Journal of Logistics, Informatics and Service Science*, vol. 5, no. 1, pp. 1–16, 2018.
- [4] P. Richard and S. Corey, "Negotiating liability allocation terms: risk, indemnity and IP," *Contract Management*, vol. 47, no. 11, pp. 42–53, 2007.
- [5] C. John and S. Robert, "Management of infringement risk of IP assets," *IP & Technology Law Journal*, vol. 20, no. 12, pp. 7–10, 2008.
- [6] P. C. Liu, "Who will pay for the "pain" of IP pledge financing? Also on the construction of multi-party participation system of IP pledge financing," *Scientific Research (New York)*, vol. 4, pp. 521–525, 2011.
- [7] H. F. He and C. H. Liu, "Risk early warning evaluation of IP pledge financing in Guangdong Province based on fuzzy comprehensive analysis method," *Research on science and technology management*, vol. 14, pp. 151–159, 2013.
- [8] C. Liu, Y. Zhang, and X. Z. Bao, "Dynamic crisis early warning of patent-backed loan for SMEs," *Science & Technology Progress and Policy*, vol. 35, no. 15, pp. 132–137, 2018.
- [9] H. F. Zhang, "Discussion on the construction of the patent pledge enterprises index system," *Studies in Science of Science*, vol. 35, no. 7, pp. 1026–1031, 2017.
- [10] V. Zemlickienė, "Using TOPSIS method for assessing the commercial potential of biotechnologies," *Journal of System and Management Sciences*, vol. 9, no. 1, pp. 117–140, 2019.
- [11] H. Y. Li, Z. M. Yuan, and S. H. Li, "Innovative enterprises intellectual property pledge loan risk assessment," *Studies in Science of Science*, vol. 35, no. 08, pp. 1253–1263, 2017.
- [12] X. N. Yin, X. Z. Bao, and L. M. Zhu, "Risk assessment of IP pledge financing based on the perspective of financing subject," *Science and Technology Management Research*, vol. 12, pp. 125–129, 2016.
- [13] X. Z. Bao, W. J. Xie, and W. Y. Dong, "Risk assessment of IP-backed financing based on an improved interval number

- TOPSIS method,” *Science and Technology Management Research*, vol. 40, no. 4, pp. 198–205, 2020.
- [14] X. Miao and X. L. Dong, “Research on risk assessment of intellectual property pledge financing for small and medium-sized technological enterprises: based on the perspective of supply chain finance,” *Science and Technology Management Research*, vol. 40, no. 18, pp. 196–202, 2020.
- [15] J. Yoon and S. Joung, “A big data based cosmetic recommendation algorithm,” *Journal of System and Management Sciences*, vol. 10, no. 2, pp. 40–52, 2020.
- [16] T. Šarić, Đ. Vukelić, K. Šimunović et al., “Modelling and prediction of surface roughness in CNC turning process using neural networks,” *Tehnicki vjesnik-Technical Gazette*, vol. 27, no. 6, pp. 1923–1930, 2020.
- [17] V. Venkataramani and K. Chokkalingam, “Analysis and implementation of positive output super lift converter using artificial neural network with sliding mode controller,” *Studies in Informatics and Control*, vol. 29, no. 3, pp. 373–383, 2020.
- [18] L. Zeng and M. Ming, “Risk assessment of IP pledge financing of small and medium sized enterprises based on BP neural network,” *Science and Technology Management Research*, vol. 36, no. 23, pp. 164–167, 2016.
- [19] Y. Qian, J. Zeng, S. Zhang, D. Xu, and X. Wei, “Short-term traffic prediction based on genetic algorithm improved neural network,” *Tehnicki vjesnik-Technical Gazette*, vol. 27, no. 4, pp. 1270–1276, 2020.
- [20] X. L. Sun, Y. Z. Tian, and W. B. Wang, “Dynamic early warning model of enterprise financial crisis based on Kalman filter,” *Systems Management Journal*, vol. 4, pp. 408–427, 2010.
- [21] Z. Z. Zhu, “Research on financial crisis early warning from the perspective of enterprise life cycle,” *Southeast University, Nanjing, China*, 2016.
- [22] G. Fu, P. Zeng, and G. Chen, “Empirical study on early warning of enterprise financial crisis under the new normal economy,” *Financial Science*, vol. 9, pp. 88–99, 2016.
- [23] J. C. Van Home and J. M. Wachowicz, *Fundamentals of Financial Management*, Prentice Hall International, New York, 1999.
- [24] S. B. Peng and J. P. Xing, *Financial Crisis Theory of the Company*, Tsinghua University Press, Beijing, 2005.
- [25] X. L. Sun, *Research on Dynamic Early Warning Model of Financial Crisis*, Shanghai Jiaotong University Press, Shanghai, China, 2011.
- [26] L. Zhou, K. K. Lai, and J. Yen, “Empirical models based on features ranking techniques for corporate financial distress prediction,” *Computers & Mathematics with Applications*, vol. 64, no. 8, pp. 2484–2496, 2012.

## Retraction

# Retracted: A Study on the Impact of Regional Total Factor Production in Digital Economy Based on Fuzzy Hierarchical VISC Algorithm

### Computational Intelligence and Neuroscience

Received 1 August 2023; Accepted 1 August 2023; Published 2 August 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] H. Li and W. Xue, "A Study on the Impact of Regional Total Factor Production in Digital Economy Based on Fuzzy Hierarchical VISC Algorithm," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 6903836, 11 pages, 2022.

## Research Article

# A Study on the Impact of Regional Total Factor Production in Digital Economy Based on Fuzzy Hierarchical VISC Algorithm

Haixiang Li<sup>1,2</sup> and Weixian Xue<sup>1</sup> 

<sup>1</sup>School of Economics and Management, Xi'an University of Technology, Xi'an 710048, China

<sup>2</sup>School of Humanities and Management, Xi'an Traffic Engineering Institute, Xi'an 710300, China

Correspondence should be addressed to Weixian Xue; 1160511009@stu.xaut.edu.cn

Received 31 December 2021; Accepted 25 January 2022; Published 26 February 2022

Academic Editor: Daqing Gong

Copyright © 2022 Haixiang Li and Weixian Xue. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The classification inaccuracy occurs to the evaluation calculation of the total factor production impact in the traditional digital economy development area. This paper applies the fuzzy hierarchical (Visualization in Scientific Computing, VISC) algorithm to the calculation method of the total factor production impact evaluation in the digital economy development area. The quantitative recursive method is used to evaluate the ability of the data information model, to achieve the ability of controlling the acquisition of characteristic resources, and complete the classification and summary of the index parameters of the total factor production impact of the digital economy development area. Finally, the experiment proves that this calculation method is used to develop the evaluation of the impact of total factor production digital economy development area and improve the information integration and analysis capabilities, the accuracy of the evaluation of the impact of total factor production, and the efficiency of the use of digital economy development resources.

## 1. Introduction

As an important transportation infrastructure, several regions are also windows for foreign economic development, providing a guarantee for regional economic development and trade. Since the new century, due to the expansion of city scale and the continuous improvement of regional functions, the digital economy development area is not only the main link of economic development, but also a key link in the economic structure. The transformation of the business structure of the digital economy development area has also become an important research topic [1, 2]. Compared with the traditional regional business development model, the all-factor production structure can effectively utilize the resource advantages of the digital economy development area, reduce the risk of operation, and improve the regional profitability. Therefore, it has been widely used in many digital economy development areas. Digital economy development areas are difficult to get rid of traditional thinking patterns when determining all-factor production strategies.

In many cases, companies do not fully consider the total factor production strategy, and the effect is usually not significant. For example, the management system is not sound or the industry to be expanded is inappropriate, which may cause the failure of the regional total factor production strategy and will eventually drag down the main production industries in the region. How to choose an appropriate business strategy for a digital economy development area will be a serious test for regional business managers [3].

After the 1990s, digital economy promoted a large number of new industries, new formats, and new models, and the research of digital economy began to sprout. So far, the research on digital economy mainly focuses on the following aspects: first come the definition and evolution characteristics of digital economy [4]. The essence of digital economy is a special economic form of trading goods and services in a virtual way. Its development is closely related to the information and communication technology industry and accelerates the penetration and



change of the operation mode of relevant industries. Chinese scholars have been studying and deepening the connotation and characteristics of digital economy since 2008. He Xiaoyin believes that digital economy is not only a trade behavior dominated by e-commerce, but also an economic process of manufacturing and providing digital products and services; Kang Tiexiang uses a nondefinition method to describe in detail the remarkable characteristics of digital economy, such as virtuality, high permeability, and values added; Yang Xinming expounds the digital economy by analyzing the economic logic of digital economy to the transformation of traditional economy. Second, it introduces the development strategy of digital economy at home and abroad [5]. In order to protect the digital content industry, the British government promulgated the digital economy 2010 law in 2010; in order to achieve the objectives of broadband construction, government development, and Internet education, the Australian government launched the national digital economy strategy in 2011. Third, there is the calculation method of digital economy. Kang Tiexiang uses the sum of the total added value of the digital industry sector and the added value created by digital auxiliary activities to calculate the scale of the digital economy.

Some scholars have found that digital economy can promote economic growth, but few explain its correlation with technological innovation. According to the new economic growth theory, innovation is the driving force of technological progress and the source of economic growth. As an important index to measure technological innovation, total factor productivity analyzes the impact of digital economy on total factor productivity, which is of great significance to speed up the development of economy. Quality development has certain reference value [6].

This article carries an in-depth analysis of the advantages and disadvantages of the selection of total factor production strategies in the digital economy development area and at the same time comprehensively considers domestic and foreign policies, industrial structure, finance, ownership of resources, and corporate managers, etc. The effective implementation of the strategy provides theoretical support and practical testing [7].

## 2. Methodology and Algorithm

To correctly evaluate the impact of total factor production in a digital economy development area, first of all, an information sample model that restricts the impact of total factor production needs a digital economy development area to be built. Combined with the fuzzy hierarchical VISC algorithm, statistical analysis of the factor production impact is carried out. The total factor production impact constraint index parameter of the digital economy development area is a set of nonlinear time series. Construct the total factor production impact analysis and evaluation parameter of the digital economy development area [8].

$$x_n = x(t_0 + n\Delta t) = h[z(t_0 + n\Delta t)] + \omega_n. \quad (1)$$

In the formula,  $\omega_n$  is the error measuring function of evaluation. The solution vector of the analysis and evaluation in the total factor production impact of the digital economy development is calculated by the correlation fusion method, and the digital economy development analysis and evaluation is obtained, and the following conditions are met.

$$\begin{aligned} \Sigma &= \text{diag}(\delta_1, \delta_2, \dots, \delta_r), \quad \delta_i = \sqrt{\lambda_i}, \quad \forall i \neq j, \\ \bigcup_{i=1}^L S_i &= V - v_s. \end{aligned} \quad (2)$$

Let  $x_{n+1} = \mu x_n (1 - x_n)$  be the influence of total factor production in a digital economy development area, where  $(I_i)_{i \in N} = \{x_1, x_2, \dots, x_m\}$ . For the multivariate group, the analysis and evaluation statistical characteristic distribution sequence  $x(n)$  of digital economy development category, the data information flow model of the total factor production impact of the digital economy development area is built based on the previous statistical measurement values:

$$\begin{aligned} c_{1x}(\tau) &= E[x(n)] = 0, \\ c_{2x}(\tau) &= E\{x(n)x(n+\tau)\} = r(\tau) \\ c_{kx}(\tau_1, \tau_2, \dots, \tau_{k-1}) &\equiv 0, k \leq 3. \end{aligned} \quad (3)$$

When  $q=2$ , the total factor production in the digital economy development area affects the classroom assessment teacher's strength level and the digital economy development resource distribution level, to meet the  $(2+1)$  dimensional continuous functional condition. In other words, the course of the development of the digital economy needs to be analyzed and evaluated.

$$\psi_x(\omega) = \ln \Phi_x(\omega) = -\frac{1}{2}\omega^2 \sigma^2. \quad (4)$$

Exclusive analysis appraisal data information flow model of digital economic development built provides an accurate data input basis for the analysis and evaluation of the total factor production and constructs a set of scalar sampling sequence components.

The control objective function for constructing the total factor production impact forecast estimation of the digital economy development area is

$$\begin{aligned} \max_{x_{a,b,d,p}} & \sum_{a \in A} \sum_{b \in B} \sum_{d \in D} \sum_{p \in P} x_{a,b,d,p} V_p, \\ \text{s.t.} & \sum_{a \in A} \sum_{d \in D} \sum_{p \in P} x_{a,b,d,p} R_p^{bw} \leq K_b^{bw}(S), \quad b \in B. \end{aligned} \quad (5)$$

The gray model is used to quantitatively and recursively evaluate the level of total factor production impact capacity of digital economy development. Assuming the historical data  $\{x_i\}_{i=1}^N$  of the distribution of the total factor production impact capacity of the digital economy development, the initial value of the feature is fixed and the budgetary estimate density functional is estimated from the factor production impact prediction of the digital economy development [9].



$$u_c(t) = Kx_c(t). \quad (6)$$

The total factor production impact continuous function of the calculation model of forecast estimation system of the digital economy development area is  $u: I \times IR^d \rightarrow IR$ ; after  $k-1$  iterations,  $k \geq 1$ , the gray-scale sequence of the total factor production impact of the digital economy development area satisfies  $N(k) < L$ , and the fuzzy hierarchical VISC algorithm is adopted, to obtain the output indicator of analysis evaluation of the impact of the economic production of all factors, which is the  $K$  adjacent sample values of the distributed big data information flow:

$$P_{1j} = \sum_{d_i=KNN} \text{Sim}(x, d_i) y(d_i, C_j). \quad (7)$$

The method of fusion of big data information is used to construct the total factor production impact and construct the objective function of analyzing and evaluating the interdomain classification of the distributed data information flow. That is, the objective function of a big data cluster is

$$J_m(U, V) = \sum_{k=1}^n \sum_{i=1}^c \mu_{ik}^m (d_{ik})^2. \quad (8)$$

The researched digital economy development curriculum is analyzed, to quantitatively analyze and evaluate the exponential correlation distribution sequence  $\{x_n\}_{n=1}^N$ , find an excellent method of  $K$  value, and obtain the quantitative recursive feature extraction result of the digital economy development analysis and evaluation:

$$x_n = a_0 + \sum_{i=1}^{M_{AR}} a_i x_{n-i} + \sum_{j=0}^{M_{MA}} b_j \eta_{n-j}. \quad (9)$$

In the formula,  $a_0$  is the sampling range of the evaluation of the impact of total factor production in the initial stage of digital economy development;  $x_{n-i}$  is scalar time series;  $b_j$  is the vibration attenuation value of the digital economy development arrangement and analysis evaluation.

Relevant data about the development of the digital economy is listed gradually in the classroom. Mainly, time, class, curriculum, classroom, teacher, and other related factors belong to the optimization decision problem of multiple factors. The fuzzy hierarchical VISC algorithm is a constant weighted evaluation, and the constant easily leads to the unfairness of the evaluation. In the course importance assessment, a part of the component must be activated, and that weight increases with the increase of the factor scores. The method of optimizing the weight vector with the state vector of each indicator  $(x_1, x_2, \dots, x_m)W = (\omega_1, \omega_2, \dots, \omega_m)$  forms the concept of incentive variable weight. The equilibrium function is integrated into variable weights, and the new variable weight formula is obtained:

$$\omega_i(x_1, x_1, \dots, x_m) = \frac{\omega_1 x_1^{a-1}}{\sum_{i=1}^m \omega_i x_i^{a-1}}. \quad (10)$$

When  $a > 1$ ,  $\omega_i(x_1, x_2, \dots, x_m)$  is the incentive variable weight. During the process of comprehensive evaluation, if an index value is very large, its comprehensive evaluation value will increase rapidly; that is,  $x_i$  will increase, and  $\omega_i$  will become larger [10].

Construct a constraint of the influence of total factor production in the digital economy development area and use the fuzzy hierarchical VISC algorithm to analyze of the influence of the total factor production in the digital economy development area. The least squares problem is to find the consistent estimate of the resource constraint vector  $\beta$  of the total factor production in the digital economy development area, so that  $\|Y - X\beta\|$  reaches the minimum, where  $\|\cdot\|$  is the F-norm in the European norm, and the entropy feature extraction value of the total factor production impact constraint feature information in the digital economy development area obtained is

$$P_{\text{loss}} = 1 - \frac{1 - p_0}{\rho} = \frac{p_0 + \rho - 1}{\rho} = \sum_{n=1}^N p_{K,n}. \quad (11)$$

Given that  $d_i$  is the disturbance feature vector for the analysis and evaluation of the total factor production impact, the estimation formula of the total factor production impact of the digital economy development area is transformed into the least square solution:

$$\begin{aligned} z(t) &= x(t) + iy(t) \\ &= a(t)e^{i\theta(t)} + n(t), \end{aligned} \quad (12)$$

where  $x(t)$  is the real part of the time series for evaluating the distribution of big data;  $y(t)$  is the imaginary part of the total factor production impact constraint index sequence for the digital economy development area.

The surrogate data method is used to randomize the amplitude of the impact of factor production in the development of the digital economy. It is also possible to perturb the empirical distribution data of the functional analysis evaluation of the  $k$ -th digital economy development and obtain the  $k$ -th subgroup, which represents the utilization rate of the resource distribution of the digital economy.

$$U_{\text{util}} = \gamma \bar{X}. \quad (13)$$

Constructing a hierarchical tree, big data analysis methods are used to establish the main component characteristics of the out-of-time analysis evaluation of the total factor production of the digital economy development, and the ambiguous close filling method is used to solve the similarity of the resource distribution of the digital economy.

$$\text{Sim}_1(d_i, d_{1j}) = \frac{\sum_{k=1}^M W_{ik} \times W_{1jk}}{\sqrt{\sum_{k=1}^M W_{ik}^2} \cdot \sqrt{\sum_{k=1}^M W_{1jk}^2}}, \quad (14)$$

where  $d_i$  is the prior distribution feature vector of the total factor production impact of the digital economy development area;  $d_{1j}$  is the  $K$ -means clustering center vector of the first-level big data.

Combining the fusion method of linear correlation characteristics, the clustering and integration of the index parameters of the total factor production impact assessment of the digital economy development is realized, and the fusion formula of the output digital economy development resource information is as follows.

$$P(w|x) = P(x|w)/P(x). \quad (15)$$

If the quantitative recursive feature is  $(N(i) \bmod L) < m$ , the probability density feature of total factor production affecting resource distribution is  $p(i) = \lfloor N(i)/L \rfloor$ , the big data stream of the digital economy development area total factor production influence  $X(i)$  is divided into  $p(i)$  sub-matrices  $X_{ij}$  whose size is  $N_{ij} \times m$ , and, through index parameter clustering and integration, the corresponding total factor production impact resource allocation plan is compiled. As a result, the optimization of the influence of all-factor production in the digital economy development area has been realized.

### 3. Analysis and Result

**3.1. Data Sampling Method.** As the most important feature of the digital economy, innovation has created a subversive business model on the “cloud network end” platform of the new generation infrastructure. Driven by Internet enterprises, the construction of a large-scale information platform has accelerated the digitization process of other industries and greatly improved the operation efficiency and business coverage radius of enterprises. The new generation infrastructure maximizes the innovation capacity and innovation potential of individual and small enterprises and reduces the information cost paid by the whole country to the greatest extent. The connection of the Internet promotes the vitality of the industry and endows the industrial innovation value. The digital economy presents connection benefits and output, driving economic growth and prosperity. At the same time, the development level of digital technology affects the digital literacy of the whole people. The digital literacy level of individuals and enterprises reflects their adaptability in the digital age. Whether massive digital information can be effectively obtained, transmitted, and utilized is a necessary condition for enterprises and individuals to improve the level of productivity. For a country, the digital literacy of the whole people determines the comprehensive quality of the country, and it is also an important factor for the country to improve innovation and development in the digital age. The drive of digital economy to regional economic growth is based on the following three aspects: first, drive the demand for innovation. Consumers in the Internet era are also data producers. Big data records commodity purchase behavior, choices, and attitudes, explores consumers’ hidden needs, and opens up a huge new consumer market by using electronic online retail. Second, drive innovation supply. As emerging production factors, cloud computing and the Internet have reshaped the market production relations, solved the current problems faced by small- and medium-sized enterprises, such as shortage of

funds and insufficient supply of labor and talents, and created emerging business systems and business supply models. Third, improve the “global village market.” Digital economy not only upgrades and transforms traditional industries such as manufacturing, agriculture, construction, and industry, but also improves and transforms service industries such as education and medical treatment with high barriers. Digital economy breaks the industrial boundary and forms a new business ecological model of cross-border industrial chain.

**3.2. Construction of Digital Economy Index of System.** Regarding the selection of digital economy indicators, refer to the indicators and measurement methods issued by the State Statistical Bureau, China Academy of Telecommunication Research of MIIT (CAICT), Tencent Research Institute, etc., and follow the high principle of digital economy development, reliability, and accuracy of indicator sources, to determine the following indicators and weights: first, there is digital infrastructure. Digital facilities are the foundation and the technical support for the development of the digital industry. Five data including Internet penetration rate, mobile phone penetration rate, Internet port access, IPv4 number, and the proportion of CN domain names to the total number of domain names are selected to measure this indicator, and the weight of each indicator is set to 20%. Secondly, there is digital ecologicalization [6, 7]. The ability of technological innovation determines the development realm of an enterprise. The high-frequency iterative technological innovation of the digital economy and the cruelly competitive business ecological model are important characteristics of the digital industry ecology. The ICT industry’s fixed investment in the total investment ratio of the whole society, the number of unicorn companies, software business revenue, software product revenue, and software integration revenue indicators are selected to measure the degree of digital ecologicalization, and the weight of each indicator is set to 20%. Thirdly, there are digital talents. The digital economy is a knowledge-based and innovative economy. High-quality talents and high-level scientific research capabilities are the infinite driving force for the sustainable development of the digital economy. In this thesis, the number of degrees awarded by the major in ordinary high schools and high schools is used as the basic data for measuring digital talents, and the weight of each index is set to 50%. Fourthly come the industry value lags behind digit. The added value of industry, the added value of the primary industry, and the added value of the tertiary industry are highly integrated with the digital industry in various fields. However, due to the lagging development of the digital economy in this period, these three indicators are selected to reflect the preliminary industrial development data of the digital economy decolonization and expropriation as a factor of the next development period. The weight of each index is equal to the weight of each level index. The weight of each index is set to 33.3%.

**3.3. Data Sources and Calculation Results.** The above indicator data comes from the State Statistical Bureau and the “Statistical Report on China’s Internet Development,” etc. The starting year of the software industry statistics is 2011, so the starting year of the data selected here is 2011. According to the data from 2011 to 2020, the software business revenue, software product revenue, system integration revenue, tertiary industry increase, and industrial value-added index are first standardized. All indicators are weighted, and the results of the digital economy index are shown in Table 1.

Measured from the digital economy index of China’s 30 provinces from 2011 to 2020, the leading regions of China’s digital economy development are mainly located in the Circum-Bohai-Sea, the Yangtze River Delta, and the Pearl River Delta. Judging from the results at the level, Guangdong Province has always occupied the top position. Based on the guidance of excellent digital infrastructure, industrial development, and policy guidance, Guangdong has become a leader. In 2020, Guangdong committed to building and improving big data foundation and social big data public service platform, etc. Beijing, Shanghai, Tianjin, and Zhejiang continue to develop digitally. These provinces are the areas with a high penetration rate of digital life (such as online car reservation, third-party online settlement, bicycle sharing, and making). The hot spots centered on Sichuan, Hubei, Fujian, Shaanxi, and Shandong have huge potential for economic development. With the provincial capital as the center, we will cooperate with neighboring cities to complement each other and form a digital ecological dialogue. In March 2020, China’s first national-level big data comprehensive pilot zone was established in Guizhou Province. In October 2020, the Inner Mongolia Autonomous Region approved China’s only national big data comprehensive test zone for infrastructure construction plans and development and also approved two span type comprehensive test zones and four regional demonstration comprehensive test zones for the Beijing-Tianjin-Hebei and Pearl River Delta. They are generally “from observation to exploration, from preliminary construction and development, partial integration, key breakthroughs to balanced development, and this road will move forward.” Large and fast-growing provinces will become a new driving force for China’s digital economy development from 2020 to 2030.

**3.4. Empirical Test.** The original data of each variable comes from the “China Science and Technology Statistical Yearbook” (2011–2020) and State Statistical Bureau. The very few missing data are supplemented by the average method. RD (registered trademark) and PG (registered trademark) are flattened according to the index based on 2011. For descriptive statistics of variables, please refer to Table 2.

The descriptive statistical results of the main indicators are shown in Table 3 and Figure 1.

**3.5. Analysis of Basic Estimation Results.** First of all, to determine whether the mixed regression model or the fixed model is to exclude the mixed regression model through F check. Second, through Hausman’s inspection, the model is

determined, and finally it is shown that the choice of fixed-effect model is scientific [11]. Here, the specific estimation results of the fixed-effect model are shown in Table 4. And the coefficients are all significant at the 1% level, which also verifies the assertion that the digital economy proposed in hypothesis 1 directly promotes regional total factor production. From the point of view of the control variables, the government intervention (gov) and the coefficient of openness to the outside world are obviously negative in the impact on the advancement of the industrial structure. This is basically consistent in the existing research, which shows that when the degree of government intervention is high, it will affect the advanced development of the industrial structure. The degree of opening to the outside world restricts the advancement of the industrial structure. This may be related to China’s long-term position at the low end of the international division of labor value chain, and most of its exports are labor-intensive products. The coefficient of marketization level (market) is significantly correct showing that the higher the degree of marketization, the more conducive it is to the realization of advanced industrial structure. Among the influencing factors of industrial structure rationalization, the coefficients of government intervention (gov), marketization level (market), and fixed asset investment level (invest) have dropped significantly, which has a significant role in promoting the rationalization of industrial structure.

In order to verify the mechanism of the digital economy on regional total factor production, this paper uses the intermediary effect model, taking human capital and scientific and technological innovation as intermediate variables, respectively, to analyze the influence mechanism of the digital economy on the advancement of the industrial structure and the rationalization of the industrial structure. The specific analysis results are shown in Table 5. Consistently with hypothesis 2, it indicates that the digital economy has an indirect impact on regional total factor production.

It can be seen from Table 5 that taking human capital as a mesomeric variable, the influence coefficient of the digital economy on human capital is positive, with a significant level of 1%, and the development of the digital economy has a positive impact on the improvement of human capital. The regression coefficients of human capital and digital economy for the advancement of the industrial structure are also significantly positive, and the RI regression coefficients for the rationalization of the industrial structure are all significantly negative; the digital economy shows that, by positively affecting human capital, it can indirectly promote the advancement of the industrial structure and the rationalization of the industrial structure. Other factors remain unchanged. Each additional unit of the comprehensive development of the digital economy will directly increase the level of industrial structure by 0.799 units. Meanwhile, the level of human capital will be increased to 2.797 units, which indirectly rises to 0.080 units. The indirect effect accounts for 9.10% of the total effect. In addition, the comprehensive development level of the digital economy is that, for every additional unit, the rationalization of the industrial structure will directly increase to 0.804 units.

TABLE 1: Some calculation results of the digital economy index from 2011 to 2020.

	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Beijing	45.7	541.1	65.1	72.3	99.8	121	1236.1	133.8	148.5	175
Tianjin	4	4.9	7.1	8.1	12.1	16.1	20.9	25.3	28.1	31.7
Hebei	3.2	4	4.8	7.5	8.8	10.2	10	10.9	13.2	15.5
Shanxi	2	2.3	2.6	2.9	3.6	4	4.3	4.3	5	5.5
Inner Mongolia	2.1	2.3	2.3	2.5	3.3	3.7	3.8	4.1	4.6	5.1
Liaoning	10.9	140.1	21.3	28.2	48.9	70.7	87.1	94.5	103.8	79.8
Jilin	4.5	5.6	6.6	6.9	9.3	10.8	10.8	12.6	14.4	17.3
Heilongjiang	3.7	4.3	4.5	4.9	5.9	6.7	6.7	7.5	8.7	10.7
Shanghai	19.3	205.1	24.8	28.7	47.5	70.1	69.8	79.9	92.6	104
Jiangsu	27.9	353.1	46.9	62.8	94.7	130.6	1537.1	177.1	204	228.8
Zhejiang	14	154.1	19.7	21.8	31.6	46.4	56.7	64.3	86.5	99.6
Anhui	3.1	3.6	4.2	4.3	5.4	6.4	6.6	8	11.2	21.2
Fujian	9.5	10.9	14.5	19.4	29.8	38.9	44.3	52.4	65.7	76.9
Jiangxi	2.8	3.2	3.3	3.4	4.2	4.8	5	5.3	6.9	8
Shandong	12	142.1	18.7	29	45.2	57.5	69.4	89.6	116.8	134.8
Henan	4.4	5.6	6.3	6.6	8.8	10.4	10.7	12.4	16.3	19
Hubei	4.8	5.8	7.7	8	10.8	13.5	23.9	33.8	37.9	46.6
Hunan	5	5.8	7.3	8.4	11.1	12.1	13.2	15.1	17.4	20
Guangdong	45.7	541.1	65.1	72.3	99.8	121	1236.1	133.8	148.5	175
Guangxi	4	4.9	7.1	8.1	12.1	16.1	20.9	25.3	28.1	31.7
Hainan	3.2	4	4.8	7.5	8.8	10.2	10	10.9	13.2	15.5
Chongqing	2	2.3	2.6	2.9	3.6	4	4.3	4.3	5	5.5
Sichuan	2.1	2.3	2.3	2.5	3.3	3.7	3.8	4.1	4.6	5.1
Guizhou	10.9	140.1	21.3	28.2	48.9	70.7	87.1	94.5	103.8	79.8
Yunnan	4.5	5.6	6.6	6.9	9.3	10.8	10.8	12.6	14.4	17.3
Shaanxi	3.7	4.3	4.5	4.9	5.9	6.7	6.7	7.5	8.7	10.7
Gansu	19.3	205.1	24.8	28.7	47.5	70.1	69.8	79.9	92.6	104
Qinghai	27.9	353.1	46.9	62.8	94.7	130.6	1537.1	177.1	204	228.8
Ningxia	14	154.1	19.7	21.8	31.6	46.4	56.7	64.3	86.5	99.6
Xinjiang	3.1	3.6	4.2	4.3	5.4	6.4	6.6	8	11.2	21.2

TABLE 2: Variable measurement indicators, symbols, and descriptive statistics.

Variable interpretation	$TEP_{it}$ Technological innovation	$DE_{it}$ Digital economy development level	$RD_{it}$ Innovative R&D investment	$PG_{it}$ Innovative R&D output	$DE_{it} * RD_{it}$ Digital economy through TFP marginal effect	$DE_{it} * PG_{it}$ Digital economy through TFP marginal effect
Mean	0.9997	25.2978	1.9491	5.5450	53.4603	191.935
Standard deviation	0.07116	40.6859	1.5922	6.9228	117.6017	496.0943
Maximum value	1.3637	227.72	13.609	63.731	1005.6	4100.6
Minimum value	0.6576	0.21	0.105	0.473	0.21	0.1750

TABLE 3: Descriptive statistics of variables.

Variable	Number of observations	Average value	Standard deviation	Minimum	Max
Indh	210	0.600	0.084	0.477	0.877
Indr	210	0.521	0.303	0.027	1.320
Dig	210	0.600	0.070	0.539	0.939
Hum	210	9.267	0.892	7.514	12.681
Innov	210	0.017	0.011	0.005	0.063
Gov	210	0.251	0.103	0.037	0.628
Open	210	0.256	0.275	0.013	1.343
Market	210	0.600	0.134	0.348	0.868
Invest	210	0.827	0.269	0.210	1.480



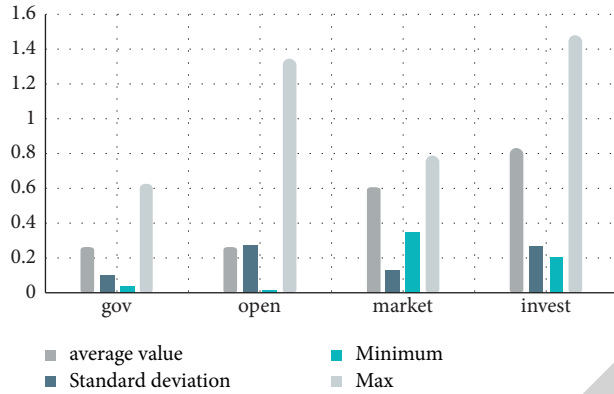


FIGURE 1: Statistical results of main variables.

TABLE 4: Benchmark regression results.

Explanatory variables	Explained variables			
	Indh	Indh	Indr	Indr
Dig	1.016*** (16.35)	0.880*** (11.83)	-1.042*** (-4.70)	-1.027*** (-3.83)
Gov		-0.193*** (-3.87)		-0.536*** (-2.98)
Open		-0.096*** (-3.09)		-0.179 (-1.60)
Market		0.177** (2.25)		-0.633** (-2.24)
Invest		-0.017 (-1.40)		-0.123*** (-2.76)
Constant term	-0.01 (-0.26)	0.053 (0.90)	1.146*** (8.60)	1.798*** (8.43)
Sample size	210	210	210	210
F value	0.599	0.653	0.11	0.218
	267.33	65.97	22.05	9.78

Note. The *t* statistic value of the estimated coefficient is shown in parentheses. \*\*\*indicates the significance level of 1%, 5%, and 10%, respectively. The following table is the same.

TABLE 5: Mesomeric effect regression results.

Variable	Explained variable					
	Hum	Indh	Indr	Innov	Indh	Indr
Mesomeric variable	Hum	Hum	Hum	Innov	Innov	Innov
Dig	2.977*** (5.35)	0.799*** (10.14)	-0.804*** (-2.81)	0.038*** (8.63)	0.776*** (8.83)	-0.736** (-2.31)
Hum		0.027** (2.74)	-0.075** (-2.07)			
Innov					2.711** (2.16)	-7.607* (-1.67)
Gov	1.206*** (3.23)	-0.226*** (-4.47)	-0.446** (-2.44)	0.008** (2.59)	-0.214*** (-4.25)	-0.478*** (-2.62)
Open	-0.330 (-1.42)	-0.087*** (-2.83)	-0.204* (-1.82)	0.005*** (2.68)	-0.109*** (-3.48)	-0.141 (-1.24)
Market	0.814 (1.39)	0.155** (2.00)	-0.572** (-2.03)	0.006 (1.24)	0.161** (2.07)	-0.589** (-2.09)
Invest	-0.033 (-0.35)	-0.017 (-1.35)	-0.126*** (-2.84)	0.001* (-1.79)	-0.021* (-1.69)	-0.113** (-2.52)
Constant term	6.801*** (15.33)	-0.131 (-1.47)	2.306*** (7.13)	-0.014*** (-3.93)	0.091 (1.48)	1.693*** (7.65)
Sample size	210	210	210	210	210	210
R <sup>2</sup>	0.347	0.668	0.237	0.401	0.662	0.231
F value	18.59	58.26	9.02	23.46	56.9	8.70

While increasing the level of human capital, it will promote the rationalization of the industrial structure, indirectly increasing by 0.233 units, with a total effect of 1.027 units. Indirect effects account for 21.71% of all effects.

When considering technological innovation as a mesomeric variable, it can be seen from Table 4 on technological innovation that the regression coefficient is significant at the 1% significance level. The regression

coefficients of technological innovation and digital economy to the advancement of the industrial structure are significantly positive. The digital economy shows that, by improving the level of technological innovation, it can indirectly promote the advancement and the rationalization of the industrial structure. Other factors remain unchanged. For each additional unit of the comprehensive development of the digital economy, the advanced industrial structure will directly increase to 0.776 units, and the level of technological innovation will be increased to 0.038 units, resulting in an indirect increase of 0.103 units in the advanced industrial structure. The effect is 0.879 units, and the indirect effect accounts for 11.72% of the total effect. In addition, for every additional unit of the comprehensive, the rationalization of the industrial structure will directly increase to 0.736 units. While the digital economy promotes the improvement of the level of technological innovation, it indirectly promotes the rationalization of the industrial structure by 0.289 units, with a total effect of 1.025 units. The indirect effect accounts for 28.20% of the total effect.

On the whole, when human capital is a mesomeric variable, the ratios of the digital economy on the advancement and rationalization of industrial structure are 1 and 4:1; when technological innovation is the intermediary variable, the ratio of the direct effect and the indirect effect of the digital economy on the advancement of the industrial structure and the rationalization of the industrial structure is 8:1 and 3:1, respectively. Whether it is based on the advancement of the industrial structure or the rationalization of the industrial structure, the direct effect of the digital economy on regional total factor production exceeds that of the digital economy, improving the indirect effect of human capital accumulation and increasing the level of technological innovation on regional total factor production [12, 13]. Therefore, from the analysis of this paper, the influence of digital economy on regional total factor production is mainly direct promotion, supplemented by indirect promotion.

**3.6. Nonlinear Analysis.** In order to investigate the nonlinear relationship between the digital economy and regional total factor production, the panel threshold model is used to test, and the level of numerical economic development is selected as the threshold variable in this paper. First, confirm the existence of the threshold variable. Empirical results are available on request (due to space limitations, no specific results are listed). There are double thresholds for the impact on the rationalization of the industrial structure of the digital economy, and each threshold is significant at a significant level of 1%, which verifies the hypothesis 3 proposed above. Based on this, the model that sets the specific threshold number is regressed, and the specific nonlinear regression results are shown in Table 6.

The reason for this phenomenon may be that if the development level of the digital economy is low, the digital economy will have a greater role in promoting labor productivity and the optimization of the proportion of the three industries to a certain extent. As the development

TABLE 6: Threshold model regression results.

Explanatory variables	Explained variables	
	Indh	Indr
dig(dig<0.777)	1.169*** (16.68)	
dig(dig≥0.777)	1.008*** (15.83)	
dig(dig<0.550)		-0.136 (-0.64)
dig(0.550≤dig≤0.563)		-0.351* (-1.68)
dig(dig>0.563)		-0.526** (-2.60)
Control variable	Control	Control
Sample size	210	210
R <sup>2</sup>	0.761	0.579
F value	92.19	33.93

level of the digital economy increases, the industrial proportional relationship is relatively complete, and the increase in labor productivity is slow. The two will not change as much as when the development level of the digital economy is low. Therefore, the role of the digital economy in promoting the advancement of the industrial structure is weakening.

When it is in the range of [0.550, 0.563], the digital economy significantly promotes the rationalization of the industrial structure. When it is greater than 0.563, the digital economy significantly affects the rationalization of the industrial structure. The promotion effect increases from 0.531 to 0.526, showing the increasing tendency of the boundary effect. As a general-purpose technology, digital technology is applied to various subindustries. The nonlinear impact of the digital economy on the rationalization of the industrial structure is due to the above characteristics. In the early stage of the development of the digital economy, the level of development of digital technology is not high, and it is mainly used in life service fields, such as medical care, education, and shopping, and there may be an insignificant phenomenon. With the rapid development of digital technology, the digital economy has gradually been applied to the field of manufacturing, which has a significant positive impact on optimizing the three industrial structures and promoting the rationalization of the industrial structure.

**3.7. Analysis of Heterogeneity.** In order to test whether on regional total factor production, this paper divides 30 provinces into three regions: east, middle, and west according to the traditional division method that is, three subsamples. Analyze the heterogeneous impact of digital economy on regional total factor production, and the regression results are shown in Table 7.

It can be seen from Table 7 that, from the perspective of the three major regions of the east, middle, and west, the digital economy has a significant role in promoting the advancement of the industrial structure and the regression coefficients are all significant at a significant level of 1%.



TABLE 7: Regression results by region.

Explanatory variables	East region		Central region		Western region	
	Indh	Indr	Indh	Indr	Indh	Indr
Gov	0.758*** (7.49)	-0.112 (-1.10)	1.487*** (10.83)	-2.114*** (-4.92)	1.218*** (6.56)	-6.056*** (-6.85)
Open	-0.481*** (-2.87)	-0.378** (-2.25)	-0.619*** (-5.83)	-1.140*** (-3.43)	-0.094* (-1.92)	-0.268 (-1.16)
Open	-0.136*** (-3.26)	-0.009 (-0.21)	0.001 (0.01)	1.173*** (3.01)	-0.182** (-2.32)	0.425 (1.14)
Market	0.113 (0.58)	0.114 (0.58)	0.123 (1.11)	-0.673* (-1.95)	0.160 (1.46)	0.829 (1.58)
Invest	-0.006 (-0.22)	0.117*** (3.97)	-0.017 (-1.08)	-0.142*** (-2.83)	-0.014 (-0.89)	-0.127* (-1.70)
Constant term	0.222 (1.36)	0.236 (1.44)	-0.212*** (-2.98)	2.348*** (10.56)	-0.105 (-1.14)	3.995*** (9.10)
Sample size	77	77	56	56	77	77
R <sup>2</sup>	0.734	0.290	0.802	0.753	0.597	0.530
F value	33.68	4.98	34.87	26.14	18.04	13.74

From the perspective of the degree of digital economy's promotion of advanced industrial structure, the impact of the digital economy has regional differences. Specifically, "the central region < western region < eastern region;" in other words, compared with the eastern region, the central and western regions can obtain more dividends. The digital economy and nonlinearity analyzed above are related to the impact of the advancement of the industrial structure. It not only promotes the advancement of the industrial structure, but also effectively narrows the gap between backward and developed regions. The impact of the digital economy on the rationalization of the industrial structure is different from the impact on the advancement of the industrial structure. The impact of the digital economy on the rationalization of the industrial structure in the eastern region is not significant. From the perspective of the central and western regions, the development of the digital economy has a significant role in promoting the rationalization of the industrial structure, and the regression coefficients are all significant at a significant level of 1%. However, the western region can get more dividends from the development of the digital economy. This proves once again that the development of the digital economy can further reduce the differences in the industrial structure between regions and promote the rationalization and development of the regional industrial structure. In view of the existence of the above economic phenomena, this paper believes that the main reasons are as follows: (1) the digital economy in the eastern region has developed rapidly and the level of development is relatively high. The industry tends to be "service-oriented," and the dividends of the digital economy that affect the production of the entire region may have been released in advance. (2) The internal gap in the development level of the digital economy in the eastern region is significantly larger than that in other regions. The comprehensive development level index at the beginning of the survey was [0.559, 0.710]. Before the final exam, this level index was in the range of [0.582, 0.939]. The development level of the digital economy with too large gap will produce the phenomenon of "overly sparse" and "overdense" elements. Moreover, because it is

not conducive to the coordinated development of the industry, the digital economy in the eastern region has an insignificant impact on the rationalization of the industrial structure.

#### 4. Discussion

Research on the impact of digital economy on economic growth and the development trend of digital economy will be done in the future. Liu Chunmei et al. studied the development process and economic characteristics of information economy (digital economy) and made an empirical regression on the quality of economic growth. The results show that information economy affects long-term economic growth by affecting the quality of economic growth. It establishes a vector autoregressive model to regress TFP and finds that information factors have a positive role in promoting TFP. The contribution of digital economy to GDP has gradually increased, bringing millions of new jobs. It is the most active field of economic development in recent years. Digital economy uses science and technology to change the current social lifestyle and industrial model and reshape the core competitiveness of the industry. Based on the above literature analysis, the following is found: firstly, scholars have enriched the meaning and characteristics of digital economy from different angles, dimensions, and directions, but it also makes it difficult to construct indicators to measure digital economy. Secondly, most scholars pay attention to the development trend and development strategy of digital economy, but the research on the impact of all-factor production of digital economy is insufficient.

In this paper, the digital economy index has a significant role in promoting technological innovation. The variable coefficients of each digital economy index are positive, indicating that the development of digital economy brings economic progress. From the perspective of interaction term coefficients, some coefficients are significantly nonzero. Therefore, when there are nonlinear interaction terms, the model can more accurately describe the current phenomenon. The negative coefficient

indicates that the smaller the number of effective invention patents of Industrial Enterprises above Designated Size, the greater the marginal effect of digital economy on total factor productivity. This shows that the degree of digital technology innovation in China is still at a low level and the current innovative industries prefer nondigital technology, resulting in the phenomenon that the marginal effect of digital technology is negative. In addition, when digital technology is integrated with real industries, there is an uneven penetration of the three industries, resulting in an unbalanced development of technology driven industries. Secondly, the impact of digital economy index on innovation in different regions is different. From the above regression results, the digital economy in the eastern coastal area plays the greatest role in promoting total factor productivity. At present, digital technology is in the early stage of integration with industry, which is manifested in the overheated development of low-end industries and the overvaluation of domestic economics, meal delivery, and other industries in the capital market. In big cities and mega cities, enterprises are facing more complex competitive markets and more serious information asymmetry, which has greatly changed the survival of the fittest mechanism. We need to pay attention not only to quality and price fluctuation, but also to marketing means and methods. The new innovation model makes the current real economy face great pressure, so the government's ability of governing the market needs to be improved.

## 5. Conclusions

In view of the research on the optimization of the total factor production impact assessment model of the digital economy development area, the digital economy development area total factor production impact method is proposed on the fuzzy hierarchical VISC algorithm proposed in this paper. The fuzzy hierarchical VISC algorithm is used to carry out an analysis of the relevant information of the total factor production impact assessment. Experiments have proved that, by using this calculation method to carry out the evaluation of the total factor production impact of the digital economy development area, the information combination analysis ability is high, the accuracy of the total factor production impact capacity evaluation is greatly improved, and the utilization efficiency of the digital economy development resources is improved.

Based on the above analysis, this paper focuses on the development of digital economy and industrial integration. There is an obvious positive correlation between the development of digital economy and regional distribution, and it has brought great demonstration effect and amplification to urban economic development. Therefore, to promote the innovation and Entrepreneurship of regional digital economy, we need not only large enterprises with leading role, but also small enterprises with strong momentum, also in addition to enterprises that dare to reform and innovate to

jointly drive the high-quality development of regional economy.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that there are no conflicts of interest.

## Acknowledgments

This research study was sponsored by the National Social Science Fund Project (no. 17BJL005) which is acknowledged.

## References

- [1] J. Tian and Y. Liu, "Research on total factor productivity measurement and influencing factors of digital economy enterprises," *Procedia Computer Science*, vol. 187, pp. 390–395, 2021.
- [2] N. I. Antipina, "Transformation of Russian business in the conditions of transition to the digital economy: sectoral and regional dimensions," *Economics of Contemporary Russia*, vol. 10, no. 1, pp. 20–22, 2018.
- [3] N. Chen and Y. Cai, "The development quality and regional characteristics of China's ict manufacturing industry in the digital economy boom: an empirical analysis based on provincial data," *Journal of Graduate School of Chinese Academy of Social Sciences*, vol. 10, no. 5, pp. 42–46, 2019.
- [4] V. S. Litvinenko, "Digital economy as a factor in the technological development of the mineral sector," *Natural Resources Research*, vol. 29, no. 3, pp. 1521–1541, 2020.
- [5] Q. Zhang, L. G. Zheng, H. Liu, and H. Cheng, "[analysis on the coordinated development of ecology-economy-society in coal resource cities: a case study of huainan, China]," *Ying yong sheng tai xue bao = The journal of applied ecology*, vol. 30, no. 12, pp. 4313–4322, 2019.
- [6] L. Nakamura, J. Samuels, and R. Soloveichik, "Measuring the free digital economy within the gdp and productivity accounts," *Bea Working Papers*, vol. 574, no. 7, pp. 914–923, 2017.
- [7] X. Nie, T. Zhang, L. I. Zhu, and F. Feng, "Ecological civilization evaluation and coordinated development between environment, economy and society in hubei," *Journal of Shanxi Normal University (Philosophy and Social Sciences edition)*, vol. 55, no. 12, pp. 24–30, 2018.
- [8] G. Liu, Y. Liu, and C. Zhang, "Factor allocation, economic growth and unbalanced regional development in China," *The World Economy*, vol. 20, no. 1, pp. 1–20, 2018.
- [9] P. Branzova, "Development of the regional bioeconomy in Bulgaria," *Economic Science, education and the real economy: Development and interactions in the digital age*, vol. 24, no. 4, pp. 14–25, 2020.
- [10] F. Wang and H. N. University, "Evaluation of coordinated development between water resources and ecological environment of wanjiang city belt," *Pearl River*, vol. 5, no. 5, pp. 3405–3407, 2017.

## Research Article

# Seismic Random Noise Denoising Using Mini-Batch Multivariate Variational Mode Decomposition

Guoning Wu <sup>1</sup>, Guochang Liu <sup>2</sup>, Junxian Wang,<sup>1</sup> and Pingping Fan<sup>1</sup>

<sup>1</sup>College of Science, China University of Petroleum (Beijing), Beijing, China

<sup>2</sup>State Key Laboratory of Petroleum Resources and Prospecting, China University of Petroleum (Beijing), Beijing, China

Correspondence should be addressed to Guoning Wu; wuguoning@gmail.com and Guochang Liu; guochang.liu@cup.edu.cn

Received 3 December 2021; Revised 20 January 2022; Accepted 21 January 2022; Published 26 February 2022

Academic Editor: Daqing Gong

Copyright © 2022 Guoning Wu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Seismic noise attenuation plays an important role in seismic interpretation. The empirical mode decomposition, synchrosqueezing wavelet transform, variational mode decomposition, etc., are often applied trace by trace. Multivariate empirical mode decomposition, multivariate synchrosqueezing wavelet transform, and multivariate variational mode decomposition were proposed for lateral continuity consideration. Due to large input data, mini-batch multivariate variational mode decomposition is proposed in this paper. The proposed method takes advantages both of variational mode decomposition and multivariate variational mode decomposition. This proposed method firstly segments the input data into a series of smaller ones with no overlapping and then applies multivariate variational mode decomposition to these smaller ones. High frequency-domain noise is filtered through sifting. Finally, the denoised smaller ones are concatenated to form components (or intrinsic mode functions) of the input signal. Synthetic and field data experiments validate the proposed method with different batch sizes and achieve higher signal-to-noise ratio than the variational mode decomposition method.

## 1. Introduction

Seismic noise attenuation plays an important role in seismic interpretation [1–3]. A variety of methods have been proposed for attenuating or removing random noise in order to enhance the signal-to-noise ratio (SNR) [4–8]. The transform-based methods, such as Fourier transform [9], wavelet transform [10], curvelet transform [11], and seislet transform [12], assume that the input signal has sparse representation with predetermined base, and under the predetermined base, noise and clean signal can be separated in the transform domain [7, 13–15].

Apart from these fixed basis methods mentioned above, there are also some data-driven methods [16]. Empirical mode decomposition (EMD) [17–19] recursively decomposes an input signal into so-called intrinsic mode functions (IMF); these IMFs are amplitude and frequency modulated subsignals with slow variations. EMD is widely used for trend detection and spectrum analysis in conjunction with Hilbert transform. Lack of a rigorous theory background for

the EMD method leaves room for other decomposition methods to come. Synchrosqueezing wavelet transform (SWT) [20], a hybrid of wavelet transform and reassignment method, squeezes values of wavelet transform to its ridges in order to sharpen the time-frequency distribution. Another data-driven method is a nonparametric one called singular spectrum analysis (SSA) [21], which firstly computes the singular value decomposition of a covariance matrix derived from the input signal. After that, SSA decomposes the input signal into a sum of components with meaningful interpretations. SSA captures the basic periodicity of an input signal and is widely used in different areas [22]. Variational mode decomposition (VMD) [23] utilizes the alternative direction method of multipliers (ADMM) and non-recursively decomposes an input signal into some principal modes. Like EMD and SSA methods, the decomposed modes of the VMD method are narrow banded and compact around some center frequencies.

Denoising methods based on EMD, VMD, and SSA, except in *f-x* fashion, are often applied trace by trace; lateral

continuity is not considered [24–26]. In order to improve SNR, the multichannel spatial coherence needs to be considered. Multivariate empirical mode decomposition (MEMD) [27] and multivariate synchrosqueezing wavelet transform (MSWT) [28], as extensions of EMD and SWT, have been proposed to try to separate multivariate modes of faster oscillations from slower ones as a whole. Recently, multivariate variational mode decomposition (MVMD) [29], an extension of VMD, has emerged to seek a collection of multivariate modulated components with minimum collective bandwidth and full signal reconstruction property. These extended methods have been used in wide areas with multivariate data analysis, such as EEG and ECG applications [30–32]. With their effectiveness of these multivariate methods, they usually accompanied with high computational complexity due to large input data. Furthermore, it is a tricky problem for the parameters' setting. In view of these situations above, mini-batch multivariate variational mode decomposition (MB-MVMD) is proposed in this paper. The proposed method firstly segments the input data into a number of batches of fixed size with no overlapping. After that, it applies VMD or MVMD for the segmented data, depending on the input data which are segmented trace by trace or not. Noise on high-frequency domain is filtered through the sifting process. Finally, the decomposed data are concatenated to form components (or IMFs) of the input signal. The proposed method has the following advantages:

- (i) VMD and MVMD are two special cases of the proposed method.
- (ii) Initial parameters can be set differently depending on different batch sizes.
- (iii) Lateral continuity is considered if the data are not segmented trace by trace.
- (iv) Instead of directly decomposing the input data as a whole, the MB-MVMD method segments the input signal into a series of smaller ones. Solutions to the smaller ones are then combined to give a solution to the original problem. This divide-and-conquer technique, therefore, promotes the computing efficiency.

## 2. From Univariate to Mini-Batch Multivariate Variational Mode Decomposition

*2.1. Univariate Variational Mode Decomposition.* The univariate variational mode decomposition seeks  $K$  number of intrinsic mode functions  $u_k(t)$  such that

$$x(t) = \sum_{k=1}^K u_k(t), \quad (1)$$

where  $u_k(t) = a_k(t)\cos\phi_k(t)$ . These modes  $u_k$  are chosen to minimize the bandwidths sum and fully reconstruct the input signal  $x(t)$ ; these can be mathematically written as [23]

$$\begin{aligned} \min_{\{u_k\}, \{\omega_k\}} & \sum_{k=1}^K \left\| \partial_t [u_k^+(t)e^{-j\omega_k t}] \right\|_2^2, \\ \text{s.t.} & \sum_{k=1}^K u_k(t) = x(t), \end{aligned} \quad (2)$$

where  $\{\omega_k\}$  denotes the center frequency and  $u_k^+(t)$  denotes the analytic signal corresponding to  $u_k(t)$ :

$$u_k^+(t) = u_k(t) + j\mathcal{H}u_k(t), \quad (3)$$

where  $\mathcal{H}$  denotes the Hilbert transform:

$$\mathcal{H}u(t) = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{u(x)}{t-x} dx. \quad (4)$$

Equation (2) uses frequency modulation and Wiener filtering techniques.

For gratified solutions, two constraints are added to optimization (2) to form a Lagrangian problem:

$$\begin{aligned} \mathcal{L}(\{u_k\}, \{\omega_k\}, \lambda) &= \frac{\alpha}{2} \sum_{k=1}^K \left\| \partial_t [u_k^+(t)e^{-j\omega_k t}] \right\|_2^2 \\ &+ \left\| x(t) - \sum_{k=1}^K u_k(t) \right\|_2^2 + \left\langle \lambda(t), x(t) - \sum_{k=1}^K u_k(t) \right\rangle. \end{aligned} \quad (5)$$

Alternative direction method of multipliers (ADMM) [23], summarized in algorithm (1), is used for the solution of optimization (5) in time domain.

For computational simplicity, the ADMM algorithm for VMD in time domain is transformed to frequency domain and is summarized in algorithm (2).

VMD nonrecursively decomposes input signal into modes with compacted bandwidths and limited amplitudes' variations. Since VMD decomposes input signal trace by trace, lateral continuity is not considered.

*2.2. Multivariate Variational Mode Decomposition.* Suppose the input signal consists of  $M$  channels, that is,  $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_M(t)]$ . As an extension of the VMD method, we try to find  $K$  multivariate modulated components  $\{\mathbf{u}_k(t)\}_{k=1}^K$  that will fully construct the input signal  $\mathbf{x}(t)$ :

$$\mathbf{x}(t) = \sum_{k=1}^K \mathbf{u}_k(t), \quad (6)$$

where the  $k$ th multivariate component  $\mathbf{u}_k(t)$  is a vector with  $M$  components:

$$\mathbf{u}_k(t) = \begin{bmatrix} u_{k,1}(t) \\ u_{k,2}(t) \\ \vdots \\ u_{k,M}(t) \end{bmatrix} = \begin{bmatrix} a_{k,1}(t)\cos(\phi_{k,1}(t)) \\ a_{k,2}(t)\cos(\phi_{k,2}(t)) \\ \vdots \\ a_{k,M}(t)\cos(\phi_{k,M}(t)) \end{bmatrix}. \quad (7)$$

Let  $\mathbf{u}_k^+(t)$  denote the Hilbert transform of  $\mathbf{u}_k(t)$ :

```

Initialization:  $u_k^0, \omega_k^0, \lambda^0$ , maxiter, eps,  $n = 0$ 
while  $n < \text{maxiter}$  and  $\text{tol} < \text{eps}$  do
  for  $k = 1: K$  do
     $u_k^{n+1} := \arg \min_{u_k} \mathcal{L}(\{u_{i < k}^{n+1}\}, u_k, \{u_{i > k}^n\}, \{\omega_i^n\}, \lambda^n)$ 
  end
  for  $k = 1: K$  do
     $\omega_k^{n+1} := \arg \min_{\omega_k} \mathcal{L}(\{u_i^{n+1}\}, \{\omega_{i < k}^{n+1}\}, \omega_k, \{\omega_{i > k}^n\}, \lambda^n)$ 
  end
   $\lambda^{n+1} := \lambda^n + \tau(x - \sum_k u_k^{n+1})$ 
   $n := n + 1$ 
end

```

ALGORITHM 1: ADMM for VMD in time domain.

```

Initialization:  $\hat{u}_k^0, \omega_k^0, \hat{\lambda}^0$ , maxiter, eps,  $n = 0$ 
while  $n < \text{maxiter}$  and  $\text{tol} < \text{eps}$  do
  for  $k = 1: K$  do
     $\hat{x}(\omega) - \sum \hat{u}_i^{n+1}(\omega)$ 
     $\hat{u}_k^{n+1}(\omega) := -\sum_{i < k} \hat{u}_i^n(\omega) + (\hat{\lambda}^n(\omega)/2)^{1/2} + 2\alpha(\omega - \omega_k^n)^2$ 
  end
  for  $k = 1: K$  do
     $\omega_k^{n+1} := \int_0^\infty \omega |\hat{u}_k^{n+1}(\omega)|^2 d\omega / \int_0^\infty |\hat{u}_k^{n+1}(\omega)|^2 d\omega$ 
  end
   $\hat{\lambda}^{n+1} := \hat{\lambda}^n + \tau(\hat{x} - \sum_k \hat{u}_k^{n+1}(\omega))$ 
   $n := n + 1$ 
end

```

ALGORITHM 2: ADMM for VMD in frequency domain.

$$\mathbf{u}_k^+(t) = \begin{bmatrix} u_{k,1}^+(t) \\ u_{k,2}^+(t) \\ \vdots \\ u_{k,M}^+(t) \end{bmatrix} = \begin{bmatrix} u_{k,1}(t) + j\mathcal{H}u_{k,1}(t) \\ u_{k,2}(t) + j\mathcal{H}u_{k,2}(t) \\ \vdots \\ u_{k,M}(t) + j\mathcal{H}u_{k,M}(t) \end{bmatrix} = \begin{bmatrix} a_{k,1}(t)\exp(j\phi_{k,1}(t)) \\ a_{k,2}(t)\exp(j\phi_{k,2}(t)) \\ \vdots \\ a_{k,M}(t)\exp(j\phi_{k,M}(t)) \end{bmatrix}. \quad (8)$$

We modulate the  $k$ th multivariate component  $\mathbf{u}_k(t)$  by  $\omega_k$ ; corresponding to equation (2), the constrained optimization problem for MVMD is

$$\begin{aligned} & \arg \min_{\{u_{k,i}\}, \{\omega_k\}} \sum_{k=1}^K \sum_{i=1}^M \left\| \partial_t [u_{k,i}^+(t) e^{-j\omega_k t}] \right\|_2^2, \\ & \text{s.t.} \quad \sum_{k=1}^K u_{k,i}(t) = x_i(t), \quad i = 1, 2, \dots, M. \end{aligned} \quad (9)$$

The Lagrangian function with added two constraints is

$$\begin{aligned} \mathcal{L}(\{u_{k,i}\}, \{\omega_k\}, \lambda_i) &= \frac{\alpha}{2} \sum_{k=1}^K \sum_{i=1}^M \left\| \partial_t [u_{k,i}^+(t) e^{-j\omega_k t}] \right\|_2^2 \\ &+ \sum_{i=1}^M \left\| x_i(t) - \sum_{k=1}^K u_{k,i}(t) \right\|_2^2 + \sum_{i=1}^M \left\langle \lambda_i(t), x_i(t) \right. \\ &\left. - \sum_{k=1}^K u_{k,i}(t) \right\rangle. \end{aligned} \quad (10)$$

Just like algorithm (1), ADMM algorithm [29] is used for the solution of equation (10) in time domain and is summarized in algorithm (3).

ADMM algorithm [29] for MVMD in frequency domain is summarized in algorithm (4); it is simpler than MVMD in time domain.

MVMD takes multivariate input signal as a whole and tries to seek  $K$  number of multivariate components from the input signal with minimum sum of bandwidths. With big data input, ADMM for MVMD in frequency domain may have high computation complexity.

**2.3. Mini-Batch Multivariate Variational Mode Decomposition.** Mini-batch multivariate variational mode decomposition (MB-MVMD) takes the advantages of both VMD and MVMD methods. MB-MVMD segments the input data into a series of smaller ones with no overlapping; after the segmentation, these smaller ones are decomposed using ADMM method just as the MVMD or VMD. This

```

Initialization:  $\{u_{k,i}^0\}, \{\omega_k^0\}, \lambda_i^0$ , maxiter, eps,  $n=0$ 
while  $n < \text{maxiter}$  and  $\text{tol} < \text{eps}$  do
  for  $k=1: K$  do
    for  $i=1: M$  do
       $u_{k,i}^{n+1} := \arg \min_{u_{k,i}} \mathcal{L} \left( \{u_{j < k,i}^{n+1}\}, u_{k,i}, \{u_{j > k,i}^n\}, \{\omega_k^n\}, \lambda_i^n \right)$ 
    end
  end
  for  $k=1: K$  do
     $\omega_k^{n+1} := \arg \min_{\omega_k} \mathcal{L} \left( \{u_{j,i}^{n+1}\}, \{\omega_{j < k}^{n+1}\}, \omega_k, \{u_{j > k}^n\}, \lambda_i^n \right)$ 
  end
  for  $i=1: M$  do
     $\lambda_i^{n+1} = \lambda_i^n + \tau(x_i - \sum_k u_{k,i}^{n+1})$ 
  end
   $n = n + 1$ 
end

```

ALGORITHM 3: ADMM for MVMD in time domain.

```

Initialization:  $\{\hat{u}_{k,i}^0\}, \{\omega_k^0\}, \hat{\lambda}_i^0$ , maxiter, eps,  $n=0$ 
while  $n < \text{maxiter}$  and  $\text{tol} < \text{eps}$  do
  for  $k=1: K$  do
    for  $i=1: M$  do
       $\hat{x}_i(\omega) = \sum_{j,i} \hat{u}_{j,i}^{n+1}(\omega)$ 
       $\hat{u}_{k,i}^{n+1} := -\sum_{j>k} \hat{u}_{j,i}^n(\omega) + (\hat{\lambda}_i^n(\omega)/2) / 1 + 2\alpha(\omega - \omega_k^n)^2$ 
    end
  end
  for  $k=1: K$  do
     $\omega_k^{n+1} := \sum_{i=1}^M \int_0^\infty \omega |\hat{u}_{k,i}^{n+1}(\omega)|^2 d\omega / \sum_{i=1}^M \int_0^\infty |\hat{u}_{k,i}^{n+1}(\omega)|^2 d\omega$ 
  end
  for  $i=1: M$  do
     $\hat{\lambda}_i^{n+1} := \hat{\lambda}_i^n + \tau(\hat{x}_i(\omega) - \sum_k \hat{u}_{k,i}^{n+1}(\omega))$ 
  end
   $n = n + 1$ 
end

```

ALGORITHM 4: ADMM for MVMD in frequency domain.

```

       $N = \text{len}(\text{Data}), \text{Batchsize}, \text{BatchNum}$ 
Initialization:  $= \lfloor N/\text{Batchsize} \rfloor \{\hat{u}_{k,i}^0\}, \{\omega_k^0\}, \hat{\lambda}_i^0$ , maxiter, eps,  $n=0$ 
for  $l=1: \text{BatchNum}$  do
   $x = \text{Data}[(l-1) * \text{Batchsize} : l * \text{Batchsize}]$ 
  while  $n < \text{maxiter}$  and  $\text{tol} < \text{eps}$  do
    for  $k=1: K$  do
      for  $i=1: M$  do
         $\hat{x}_i(\omega) = \sum_{j,i} \hat{u}_{j,i}^{n+1}(\omega)$ 
         $\hat{u}_{k,i}^{n+1} := -\sum_{j>k} \hat{u}_{j,i}^n(\omega) + (\hat{\lambda}_i^n(\omega)/2) / 1 + 2\alpha(\omega - \omega_k^n)^2$ 
      end
    end
    for  $k=1: K$  do
       $\omega_k^{n+1} := \sum_{i=1}^M \int_0^\infty \omega |\hat{u}_{k,i}^{n+1}(\omega)|^2 d\omega / \sum_{i=1}^M \int_0^\infty |\hat{u}_{k,i}^{n+1}(\omega)|^2 d\omega$ 
    end
    for  $i=1: M$  do
       $\hat{\lambda}_i^{n+1} := \hat{\lambda}_i^n + \tau(\hat{x}_i(\omega) - \sum_k \hat{u}_{k,i}^{n+1}(\omega))$ 
    end
     $n = n + 1$ 
  end
end

```

ALGORITHM 5: ADMM for MB-MVMD in frequency domain.



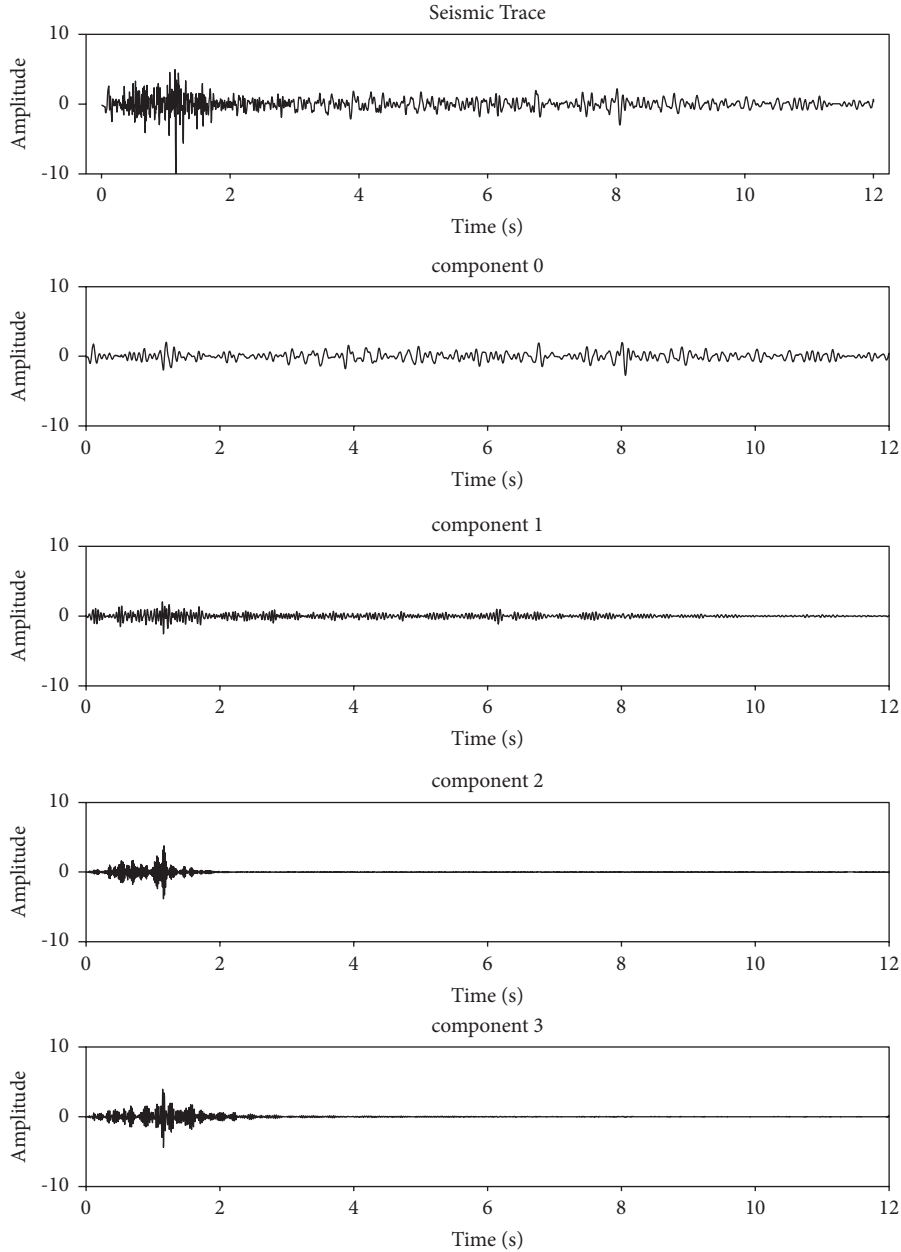


FIGURE 1: VMD decomposition. From the top to bottom are the input signal and its VMD decomposition components.

mini-batch technique not only considers lateral continuity of the input data but also promotes the computational efficiency using divided-and-conquer technique.

Suppose the input data are  $\mathbf{X}$  with  $N$  traces (or columns). We firstly set the batch size and then compute the number of batches using floor function “ $\lfloor \cdot \rfloor$ ”:

$$\text{Number of batches} = \left\lfloor \frac{N}{\text{Batch size}} \right\rfloor. \quad (11)$$

After that, we extract the  $l$ th mini-batch data from the input data  $\mathbf{X}$ :

$$\mathbf{x}_l = \mathbf{X}[(l-1) * (\text{bs}) : l * (\text{bs})], \quad (12)$$

where “bs” denotes the batch size previously determined. Following the segmentation above, we use ADMM algorithm in frequency domain to decompose the  $l$ th mini-batch data and lastly concatenate the decomposed data to form components.

The computation processes of MB-MVMD in frequency domain are summarized in algorithm (5).

### 3. Experimental Results

Figure 1 shows a single seismic trace and its four extracted components using the VMD method. These four components are narrow banded. Figure 2 shows the sum of the four

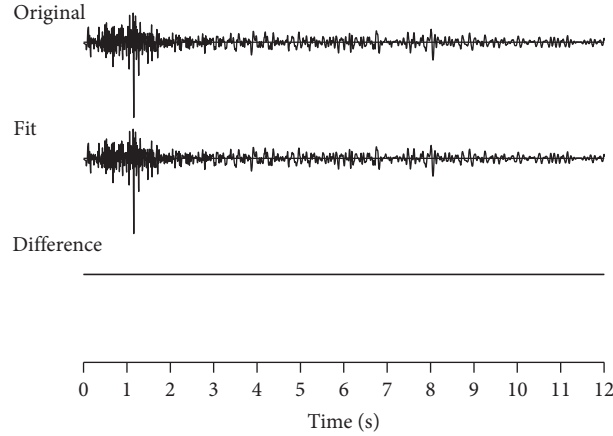


FIGURE 2: Seismic trace and its sum of VMD components: “Original” is the input seismic signal, “Fit” is the sum of VMD components in Figure 1, and “Difference” is the difference between the input signal and its VMD approximation.

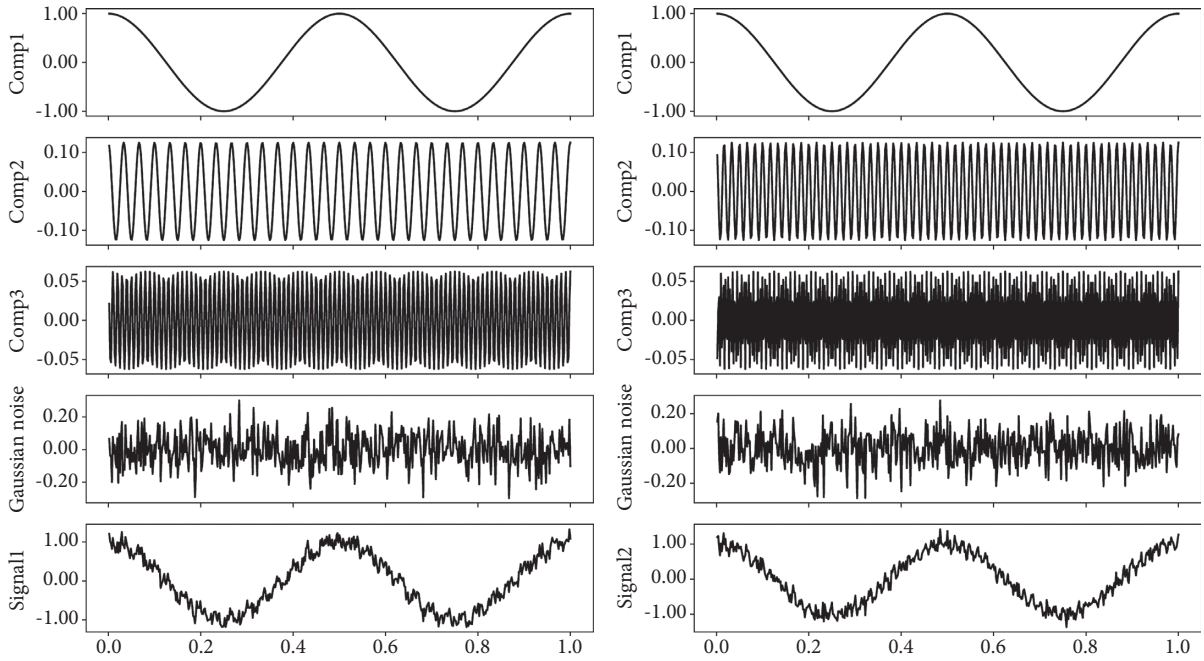


FIGURE 3: Bivariate signal: each has three components and a certain percentage of Gaussian noise. (a) The bottom of left column consists of three different tones (the top three signals in left column) and a certain percentage of Gaussian noise (the fourth signal in left column). (b) The bottom of right column also consists of three different tones and a certain percentage of Gaussian noise.

extracted components in Figure 1 and the difference between the sum and the original input single seismic trace. From these figures, we see that VMD extracts sub signals with compacted bandwidth subject to full signal reconstruction. The parameters for VMD decomposition are number of components  $K=4$  and bandwidth constraint  $\alpha = 500$ .

A bivariate signal with three different tones and a certain percentage of Gaussian noise is shown in Figure 3. We apply MVMD to this signal; the extracted components are shown in Figure 4(a). The parameters for the MVMD method are number of components  $K=4$  and bandwidth constraint  $\alpha = 500$ . For comparison, the MEMD method is also used to

decompose this bivariate signal; results are shown in Figure 4(b). From these figures, we see that MVMD uses fewer components than the MEMD method to represent the input signal.

SNR is often used as a qualitative indicator to show the effectiveness of a denoising method; it is defined as the ratio of signal power to the noise power and is often expressed in decibels:

$$\text{SNR} = 20\log_{10} \left[ \frac{A_s}{A_n} \right], \quad (13)$$

where  $A_s$  and  $A_n$  represent signal and noise powers.

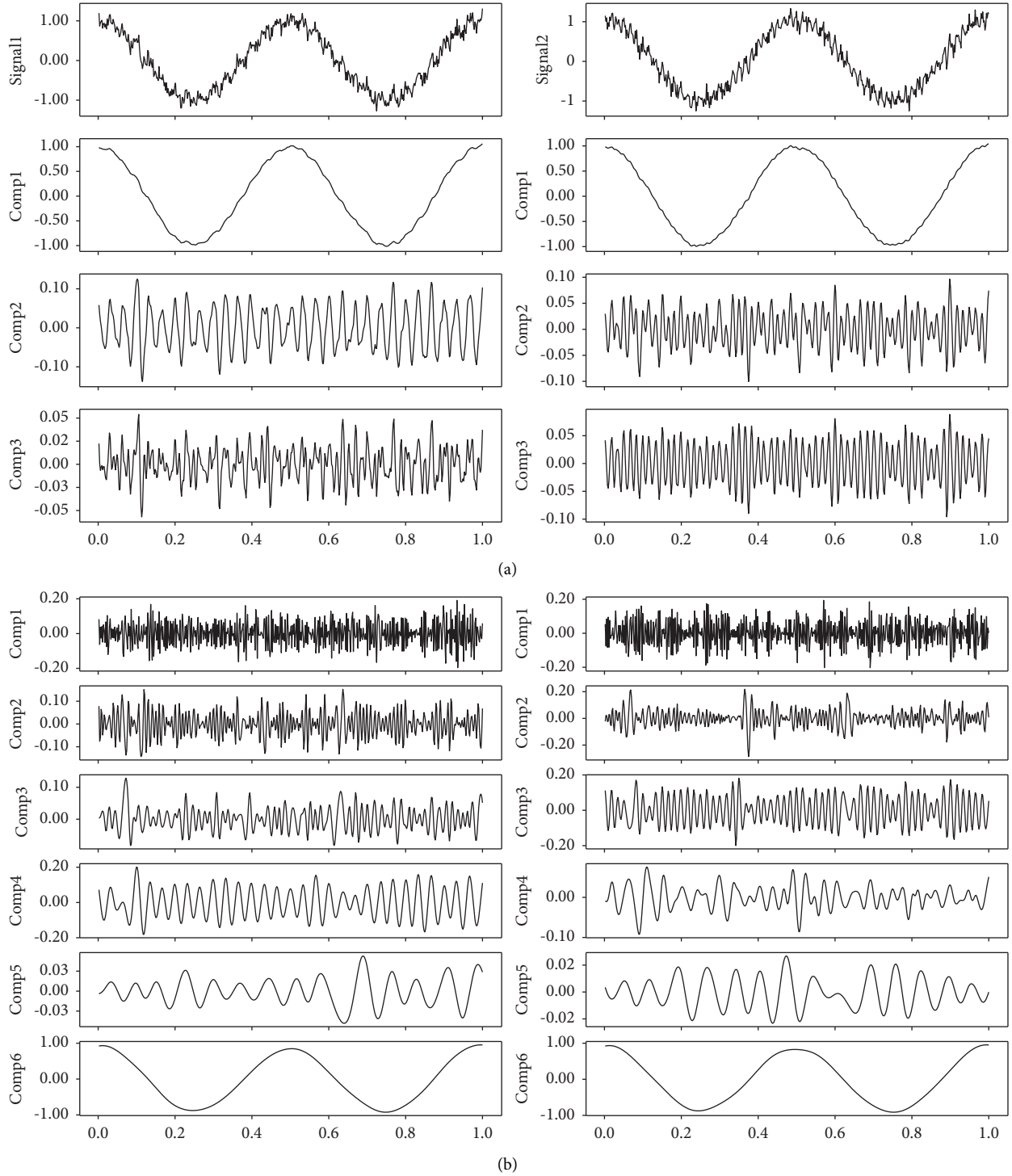


FIGURE 4: Decomposition of the bivariate signal in Figure 3. (a) MVMD: three components. (b) EMD: six components.

**3.1. Synthetic Data.** A synthetic seismic data with four linear events are used for test. The synthetic data have 256 traces, with time step  $\Delta t = 0.004$  and 512 samples. Figures 5(a), 5(b), and 5(c) are the synthetic signal, the random Gaussian noise, and the noisy data, respectively. Figure 6 shows denoised results using three different methods: FX-DECON, VMD, and MB-MVMD. Figures 6(a), 6(b), and 6(c) show the denoised result, removed noise using FX-DECON, and

the similarity between them. Figures 6(d), 6(e), and 6(f) show the denoised result, removed noise using VMD, and the similarity between them. The parameters of the VMD method are bandwidth constraint  $\alpha = 2000$  and  $\tau = 0.0$  and number of components  $K=4$ . Figures 6(g), 6(h), and 6(i) show the denoised result, removed noise using MB-MVMD of batch size eight, and the similarity between them. The parameters of the MB-MVMD method are



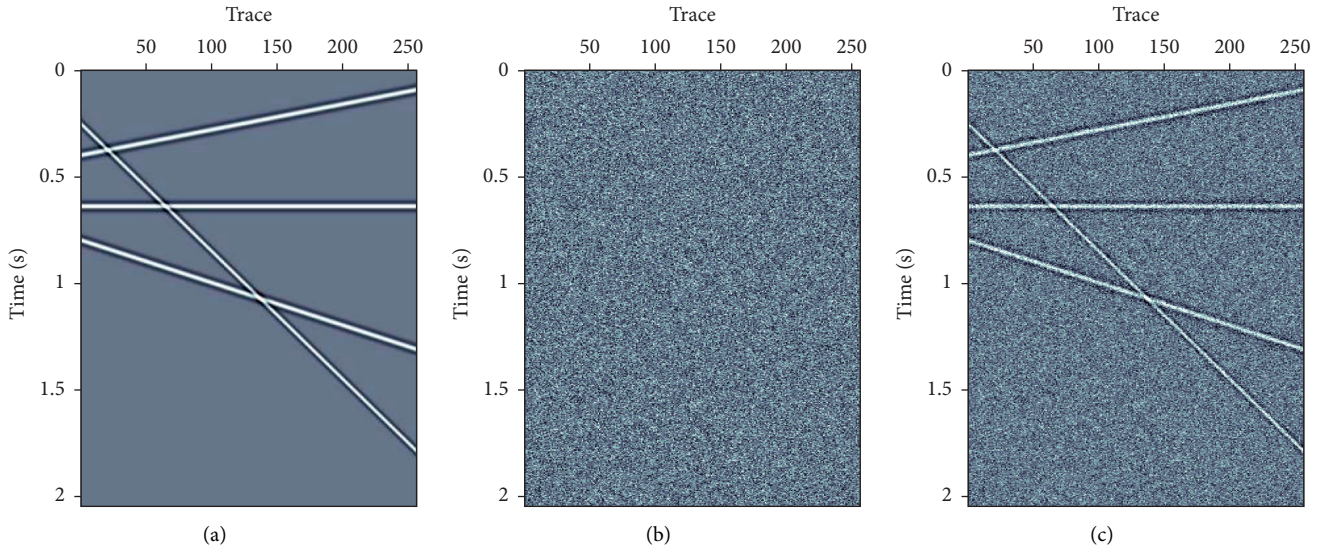


FIGURE 5: Synthetic seismic data with four linear events. (a) Clean data. (b) Gaussian random noise. (c) Data with noise added.

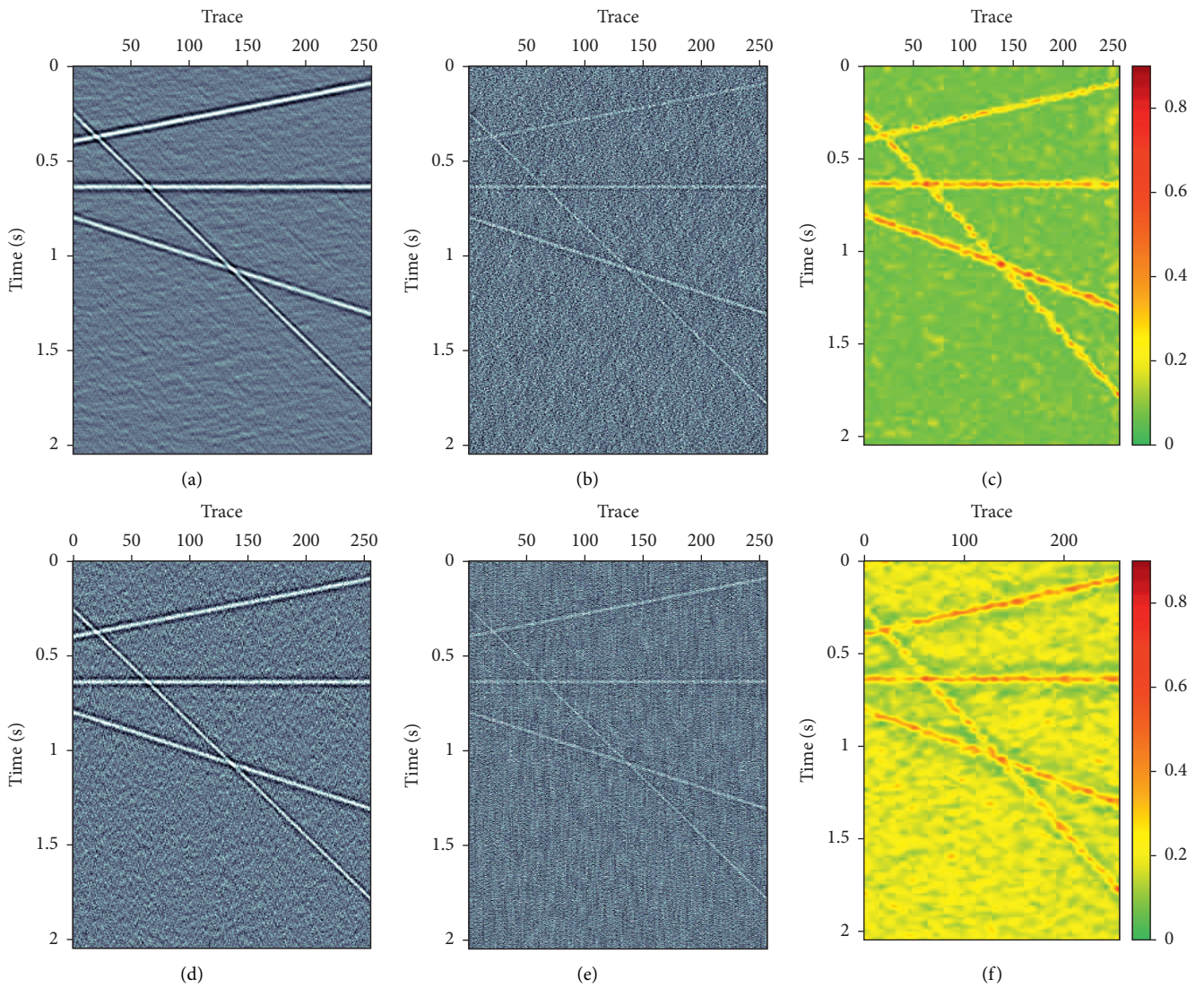


FIGURE 6: Continued.



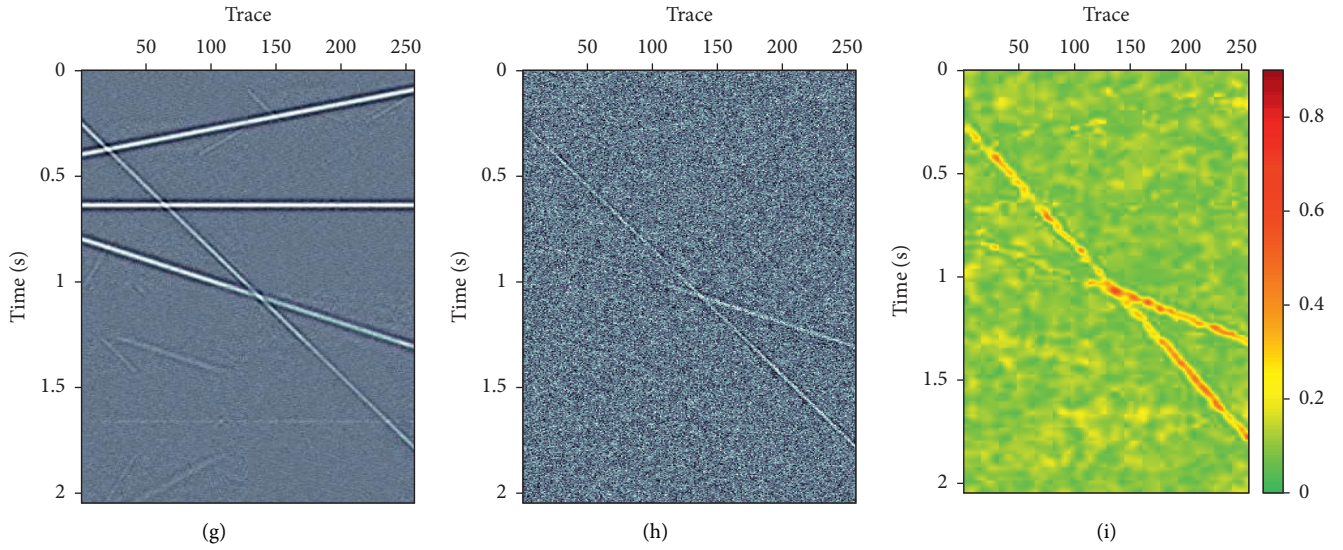


FIGURE 6: Denoise comparisons. (a) Denoised data using FX-DECON. (b) Removed noise using FX-DECON. (c) Similarity between denoised data and its removed noise using FX-DECON. (d) Denoised data using VMD. (e) Removed noise using VMD. (f) Similarity between denoised data and its removed noise using VMD. (g) Denoised data using MB-MVMD of batch size eight. (h) Removed noise using MB-MVMD of batch size eight. (i) Similarity between denoised data and its removed noise using MB-MVMD of batch size eight.

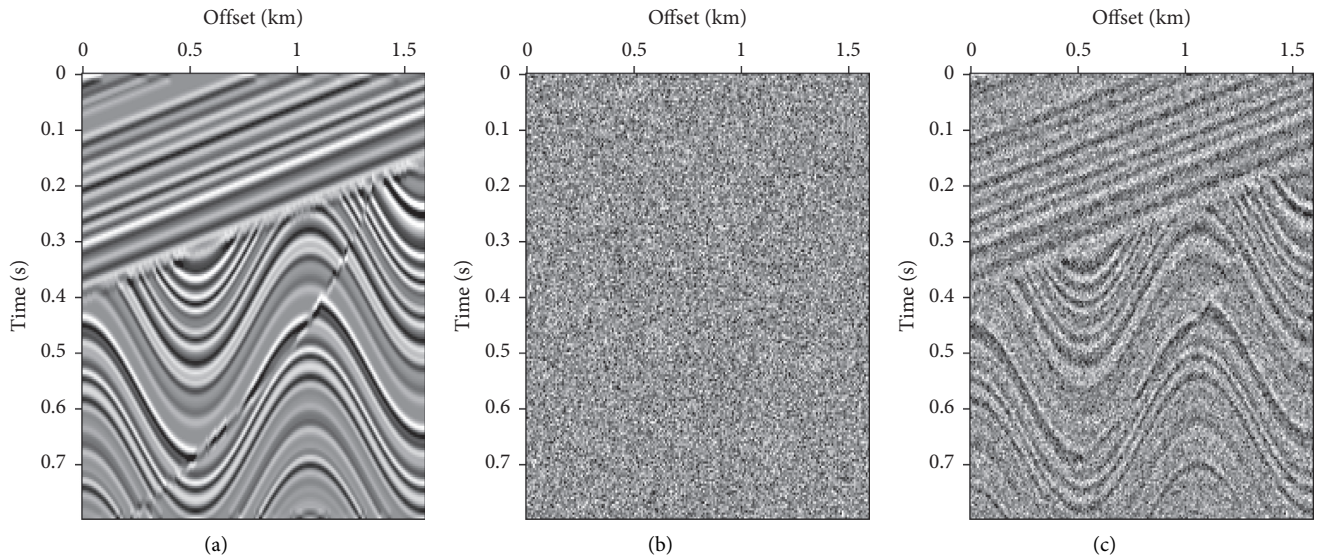


FIGURE 7: Synthetic seismic data with Gaussian noise. (a) Clean data. (b) Gaussian random noise. (c) Data with noise added.

$\alpha = 1000$ ,  $\tau = 0.5$ , and  $K = 4$ . From these figures, we can see that the denoised data of FX-DECON is better than the denoised one of VMD. The denoised data of MB-MVMD of batch size eight are the best, which can be seen from the denoised results, the removed noises, and the similarities of these three methods.

Another synthetic mode (shown in Figure 7) is used to test the proposed method. Figures 8(a), 8(b), and 8(c) are the denoised data, removed noise using FX-DECON, and similarity between them. Figures 8(d), 8(e), and 8(f) are the denoised data, removed noise using VMD, and similarity

between them. Figures 8(g), 8(h), and 8(i) are the denoised data, removed noise using MB-MVMD of batch size four, and similarity between them. The parameters of the VMD method are bandwidth constraint  $\alpha = 2000$  and  $\tau = 0.0$  and number of components  $K = 4$ . Parameters for MB-MVMD of batch size four are  $\alpha = 1000$ ,  $\tau = 0.5$ , and  $K = 4$ . The denoised result of FX-DECON is better than the denoised result of VMD; the denoised data of MB-MVMD are the best among these three denoising methods. The MB-MVMD method considers the lateral continuity of the input data; some degree of signal is removed as noise, which can be



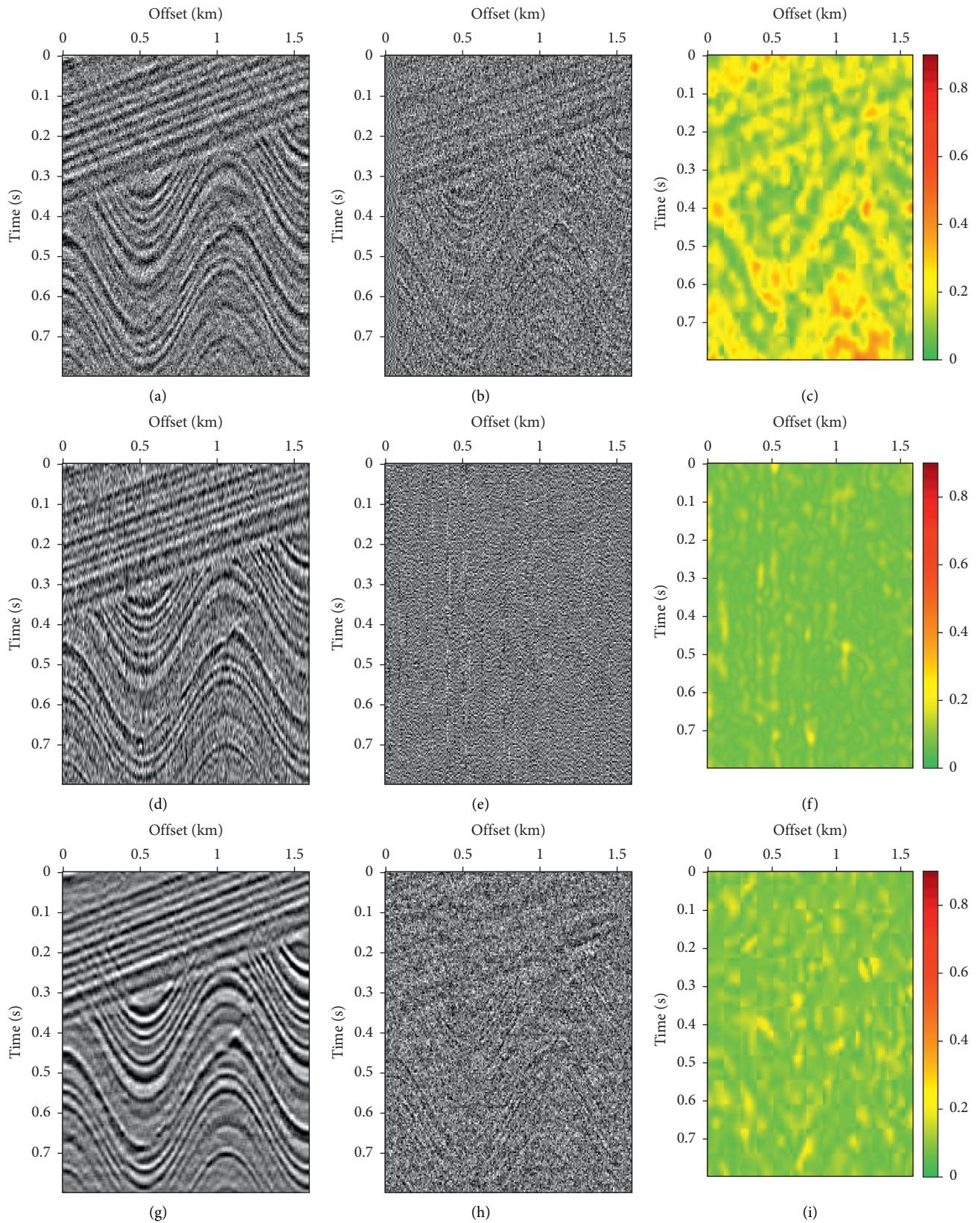


FIGURE 8: Denoise comparisons. (a) Denoised data using FX-DECON. (b) Removed noise using FX-DECON. (c) Similarity between denoised data and its removed noise using FX-DECON. (d) Denoised data using VMD. (e) Removed noise using VMD. (f) Similarity between denoised data and its removed noise using VMD. (g) Denoised data using MB-MVMD of batch size four. (h) Removed noise using MB-MVMD of batch size four. (i) Similarity between denoised data and its removed noise using MB-MVMD of batch size four.



TABLE 1: SNRs (in decibels) comparison using different batch sizes.

Batch size	2	4	8	16	32	64	128	256
Model 1	4.78	4.81	<b>6.89</b>	6.55	6.43	6.10	5.97	5.43
Model 2	5.89	<b>6.27</b>	6.12	5.83	5.74	5.61	5.57	4.98

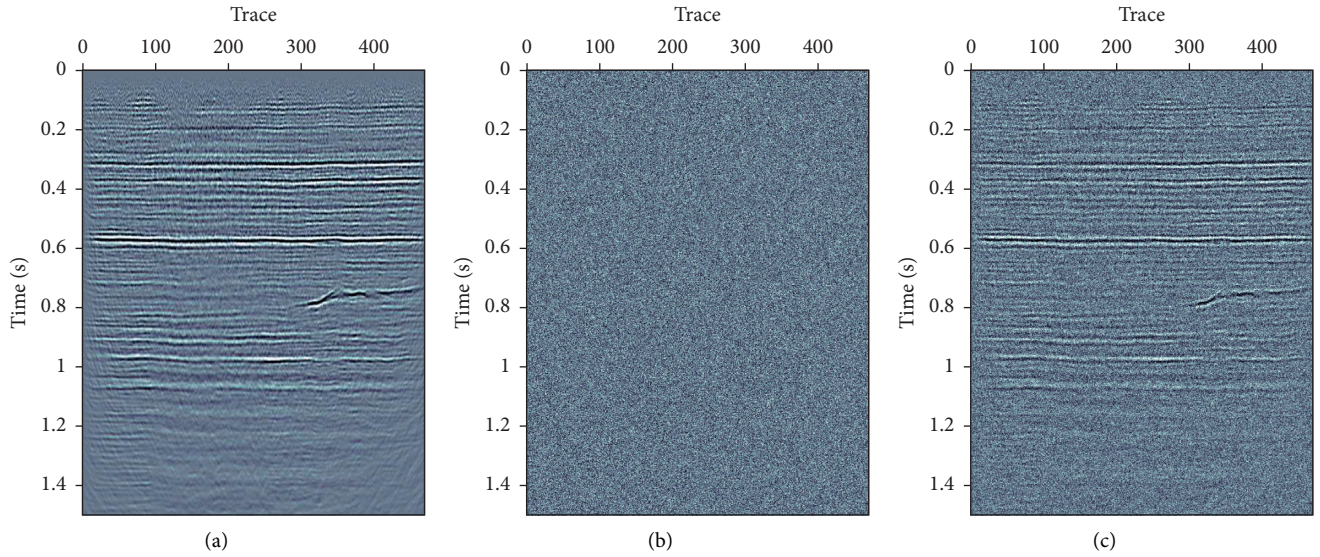


FIGURE 9: Field data with Gaussian noise. (a) Field data. (b) Gaussian random noise. (c) Field data with noise added.

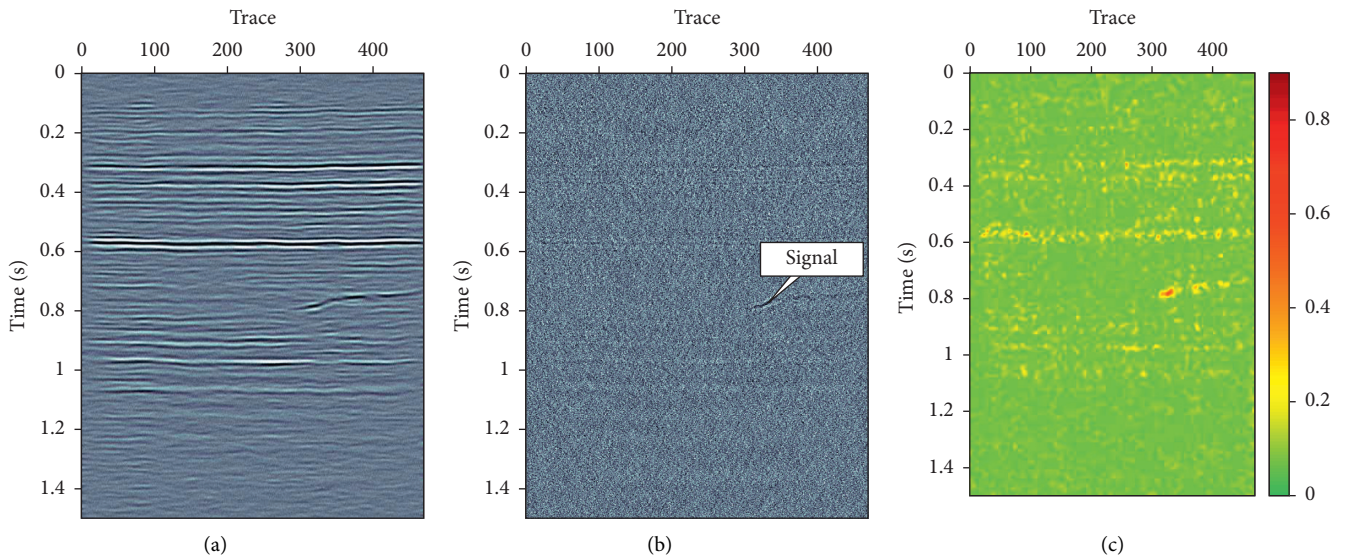


FIGURE 10: Continued.



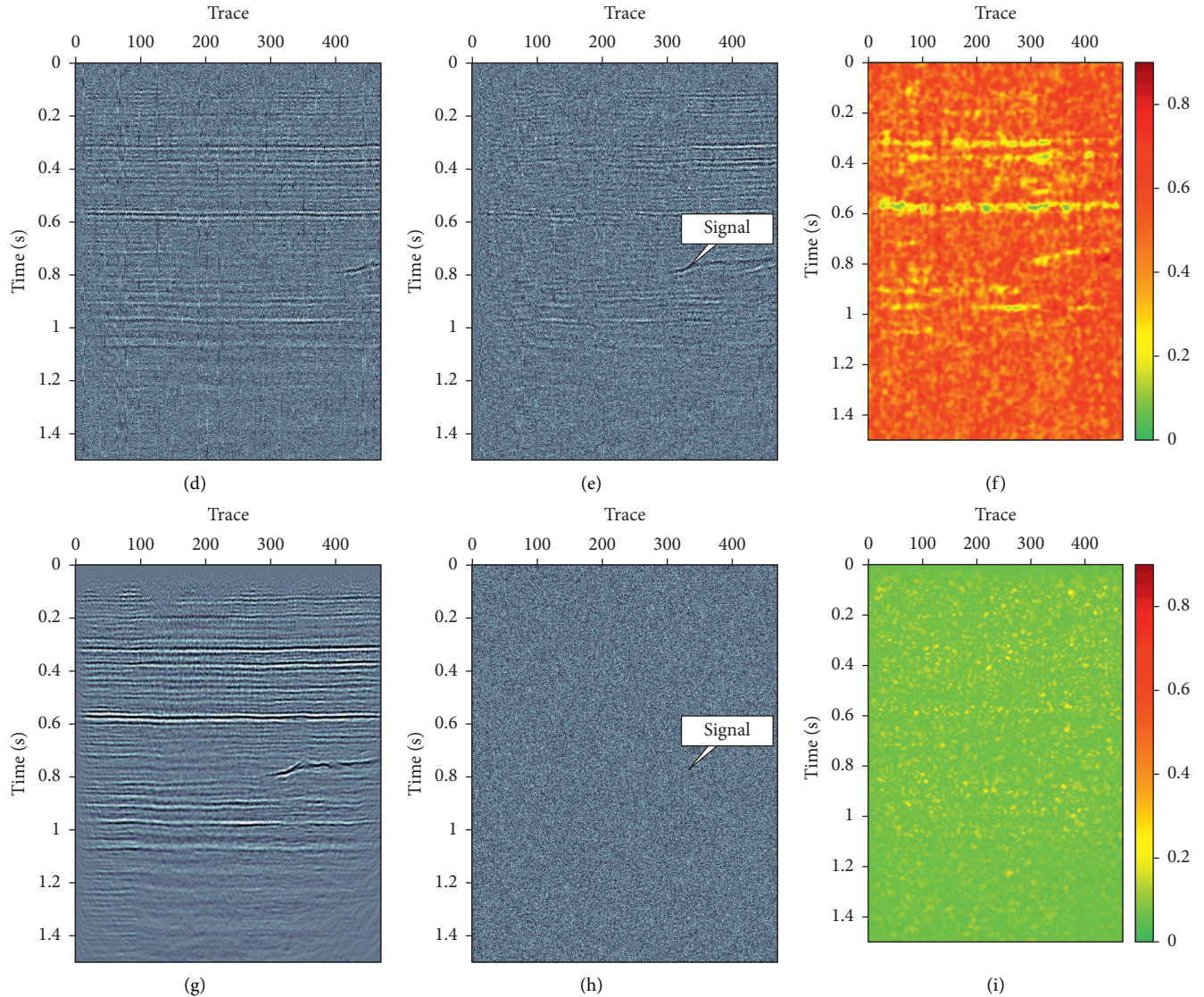


FIGURE 10: Denoise comparisons. (a) Denoised data using FX-DECON. (b) Removed noise using FX-DECON. (c) Similarity between denoised data and its removed noise using FX-DECON. (d) Denoised data using VMD. (e) Removed noise using VMD. (f) Similarity between denoised data and its removed noise using VMD. (g) Denoised data using MB-MVMD of batch size four. (h) Removed noise using MB-MVMD of batch size four. (i) Similarity between denoised data and its removed noise using MB-MVMD of batch size four.

seen from the removed noise data. Similarity of MB-MVMD of batch size four also reveals the effectiveness of the proposed method in random noise attenuation.

Table 1 shows the SNRs of the above two models using different batch sizes (Model 1 is the linear synthetic model and Model 2 is the synthetic model of Figure 7). The results show that the SNR of batch size eight is best for model 1, and the SNR of batch size four is best for model 2.

**3.2. Field Data.** Figure 9 shows a marine data; the data have 470 traces. Figures 9(a), 9(b), and 9(c) are the clean field data, Gaussian random noise, and the noisy data, respectively.

FX-DECON, VMD, and MB-MVMD of batch size five are used to denoise the noisy data. Figures 10(a), 10(b), and

10(c) are the denoised data, the removed noise using FX-DECON, and the similarity between them. Figures 10(d), 10(e), and 10(f) are the denoised data, the removed noise using VMD, and the similarity between them. Figures 10(g), 10(h), and 10(i) are the denoised data, the removed noise using MB-MVMD of batch size five, and the similarity between them. The removed noises show that a lot of signals are removed from the data for the VMD method, which is confirmed by the similarity of the VMD method (the “Signal” boxes indicate signal is removed as noise.). The removed signal as noise is least for the MV-MVMD method, which is also confirmed by the similarity of the MV-MVMD method.

Table 2 shows the SNRs of the three models using different denoising methods. “Model 1” is the synthetic model with linear events, “Model 2” is the second model, and “Field

TABLE 2: SNRs' (in decibels) comparison using FX-DECON, VMD, and MB-MVMD.

	FX-DECON	VMD	MB-MVMD
Model 1 (linear events)	4.78	-0.94	6.89
Model 2 (synthetic data)	1.81	1.25	6.27
Field data	5.90	-1.87	7.35

Data" is the field data model. The SNRs reveal that the MV-MVMD method is the best with respect to denoising results.

#### 4. Conclusion

MB-MVMD considers the lateral continuity of the input data; it seeks a sparse representation of the input signal. This divide-and-conquer method bridges VMD and MMVD methods. The proposed method achieves better denoising results compared with the VMD method in seismic random noise denoising.

Although the proposed method has many advantages, there is still room for improvement. For example, the decomposition parameters are set manually; is there a way to automatically select the best parameters for the decomposition?

#### Data Availability

The data are not freely available due to third-party rights.

#### Conflicts of Interest

The authors declare that they have no conflicts of interest.

#### Acknowledgments

The authors would like to thank the Madagascar open software for easily turning our ideas into reality. The research was partially supported by Science Foundation of China University of Petroleum (no. 2462020YXZZ006).

#### References

- [1] Y. Chen and S. Fomel, "Random noise attenuation using local signal-and-noise orthogonalization," *Geophysics*, vol. 80, no. 6, pp. WD1-WD9, 2015.
- [2] G. Liu, X. Chen, J. Du, and K. Wu, "Random noise attenuation using f-x regularized nonstationary autoregression," *Geophysics*, vol. 77, no. 2, pp. V61-V69, 2012.
- [3] W. Liu, S. Cao, Y. Chen, and S. Zu, "An effective approach to attenuate random noise based on compressive sensing and curvelet transform," *Journal of Geophysics and Engineering*, vol. 13, no. 2, pp. 135-145, 2016b.
- [4] Y. Chen, J. Ma, and S. Fomel, "Double-sparsity dictionary for seismic noise attenuation," *Geophysics*, vol. 81, no. 2, pp. V103-V116, 2016.
- [5] G. Ebrahim, "Multichannel antileakage least-squares spectral analysis for seismic data regularization beyond aliasing," *Acta Geophysica*, vol. 67, no. 5, pp. 1895-7455, 2019.
- [6] E. Ghaderpour, W. Liao, and M. P. Lamoureux, "Antileakage least-squares spectral analysis for seismic data regularization and random noise attenuation," *Geophysics*, vol. 83, no. 3, pp. V157-V170, 2018.
- [7] Z. Liu and K. Lu, "Convolutional sparse coding for noise attenuation in seismic data," *Geophysics*, vol. 86, no. 1, pp. V23-V30, 2021a.
- [8] C. Zhang and M. van der Baan, "Complete and representative training of neural networks: a generalization study using double noise injection and natural images," *Geophysics*, vol. 86, no. 3, pp. V197-V206, 2021.
- [9] J. Allen, "Short term spectral analysis, synthesis, and modification by discrete Fourier transform," *IEEE Transactions on Acoustics, Speech, & Signal Processing*, vol. 25, no. 3, pp. 235-238, 1977.
- [10] A. Chakraborty and D. Okaya, "Frequency-time decomposition of seismic data using wavelet-based methods," *Geophysics*, vol. 60, no. 6, pp. 1906-1916, 1995.
- [11] M. S. Oliveira, M. V. C. Henriques, F. E. A. Leite, G. Corso, and L. S. Lucena, "Seismic denoising using curvelet analysis," *Physica A: Statistical Mechanics and Its Applications*, vol. 391, no. 5, pp. 2106-2110, 2012.
- [12] S. Fomel and Y. Liu, "Seislet transform and seislet frame," *Geophysics*, vol. 75, no. 3, pp. V25-V38, 2010.
- [13] L. Bin, Y. Jinghang, Z. Zhiwu et al., "Unsupervised deep learning for random noise attenuation of seismic data," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1-5, 2021.
- [14] O. M. Saad and Y. Chen, "Deep denoising autoencoder for seismic random noise attenuation," *Geophysics*, vol. 85, no. 4, pp. V367-V376, 2020.
- [15] H. Wang and Y. Chen, "Adaptive frequency-domain nonlocal means for seismic random noise attenuation," *Geophysics*, vol. 86, no. 2, pp. V143-V152, 2021.
- [16] G. Wu, S. Fomel, and Y. Chen, "Data-driven time-frequency analysis of seismic data using non-stationary prony method," *Geophysical Prospecting*, vol. 66, no. 1, pp. 85-97, 2018.
- [17] N. E. Huang, Z. Shen, S. R. Long et al., "The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903-995, 1998.
- [18] J. Li, Y. Li, Y. Li, and Z. Qian, "Downhole microseismic signal denoising via empirical wavelet transform and adaptive thresholding," *Journal of Geophysics and Engineering*, vol. 15, no. 6, pp. 2469-2480, 2018.
- [19] M. Torres, M. Colominas, G. Schlotthauer, and P. Flandrin, "A complete ensemble empirical mode decomposition with adaptive noise," in *Proceedings of the In IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4144-4147, IEEE, Prague, Czech Republic, May 2011.
- [20] I. Daubechies, J. Lu, and H.-T. Wu, "Synchrosqueezed wavelet transforms: an empirical mode decomposition-like tool," *Applied and Computational Harmonic Analysis*, vol. 30, no. 2, pp. 243-261, 2011.
- [21] D. H. Schoellhamer, "Singular spectrum analysis for time series with missing data," *Geophysical Research Letters*, vol. 28, no. 16, pp. 3187-3190, 2001.
- [22] D. Zhang, Y. Chen, W. Huang, and S. Gan, "Expression of Concern: multi-step damped multichannel singular spectrum analysis for simultaneous reconstruction and denoising of 3D

- seismic data,” *Journal of Geophysics and Engineering*, vol. 16, no. 5, 1009 pages, 2019.
- [23] K. Dragomiretskiy and D. Zosso, “Variational mode decomposition,” *IEEE Transactions on Signal Processing*, vol. 62, no. 3, pp. 531–544, 2014.
- [24] M. Bekara and M. van der Baan, “Random and coherent noise attenuation by empirical mode decomposition,” *Geophysics*, vol. 74, no. 5, pp. V89–V98, 2009.
- [25] W. Liu, S. Cao, and Y. Chen, “Applications of variational mode decomposition in seismic time-frequency analysis,” *Geophysics*, vol. 81, no. 5, pp. V365–V378, 2016a.
- [26] W. Liu, S. Cao, and Z. Wang, “Application of variational mode decomposition to seismic random noise reduction,” *Journal of Geophysics and Engineering*, vol. 14, no. 4, pp. 888–898, 2017.
- [27] N. Rehman and D. P. Mandic, “Multivariate empirical mode decomposition,” *Proceedings of the Royal Society A: Mathematical, Physical & Engineering Sciences*, vol. 466, no. 2117, pp. 1291–1302, 2010.
- [28] A. Ahrabian and D. Mandic, “A class of multivariate denoising algorithms based on synchrosqueezing,” *IEEE Transactions on Signal Processing*, vol. 63, no. 9, pp. 2196–2208, 2015.
- [29] N. u. Rehman and H. Aftab, “Multivariate variational mode decomposition,” *IEEE Transactions on Signal Processing*, vol. 67, no. 23, pp. 6039–6052, 2019.
- [30] A. Hemakom, A. Ahrabian, D. Looney, N. Rehman, and D. Mandic, “Nonuniformly sampled trivariate empirical mode decomposition,” in *Proceedings of the In 2015 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2015*, pp. 3691–3695, South Brisbane, Queensland, Australia, April 2015.
- [31] H. Huan, H. Wang, and N. Rehman, “A joint framework for multivariate signal denoising using multivariate empirical mode decomposition,” *Signal Processing*, vol. 135, pp. 263–273, 2017a.
- [32] A. Zahra, N. Kanwal, N. Ur Rehman, S. Ehsan, and K. D. McDonald-Maier, “Seizure detection from EEG signals using multivariate empirical mode decomposition,” *Computers in Biology and Medicine*, vol. 88, pp. 132–141, 2017.

## Retraction

# Retracted: Multirobot Collaborative Pursuit Target Robot by Improved MADDPG

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.





### References

- [1] X. Zhou, S. Zhou, X. Mou, and Y. He, "Multirobot Collaborative Pursuit Target Robot by Improved MADDPG," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4757394, 10 pages, 2022.



## Research Article

# Multirobot Collaborative Pursuit Target Robot by Improved MADDPG

Xiao Zhou <sup>1</sup>, Song Zhou <sup>1</sup>, Xingang Mou <sup>1</sup> and Yi He <sup>2</sup>

<sup>1</sup>School of Mechanical and Electronic Engineering, Wuhan University of Technology, Wuhan 430070, China

<sup>2</sup>Intelligent Transport Systems Research Center, Wuhan University of Technology, Wuhan 430063, China

Correspondence should be addressed to Yi He; [heyi@whut.edu.cn](mailto:heyi@whut.edu.cn)

Received 19 January 2022; Accepted 8 February 2022; Published 25 February 2022

Academic Editor: Daqing Gong

Copyright © 2022 Xiao Zhou et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Policy formulation is one of the main problems in multirobot systems, especially in multirobot pursuit-evasion scenarios, where both sparse rewards and random environment changes bring great difficulties to find better strategy. Existing multirobot decision-making methods mostly use environmental rewards to promote robots to complete the target task that cannot achieve good results. This paper proposes a multirobot pursuit method based on improved multiagent deep deterministic policy gradient (MADDPG), which solves the problem of sparse rewards in multirobot pursuit-evasion scenarios by combining the intrinsic reward and the external environment. The state similarity module based on the threshold constraint is as a part of the intrinsic reward signal output by the intrinsic curiosity module, which is used to balance overexploration and insufficient exploration, so that the agent can use the intrinsic reward more effectively to learn better strategies. The simulation experiment results show that the proposed method can improve the reward value of robots and the success rate of the pursuit task significantly. The intuitive change is obviously reflected in the real-time distance between the pursuer and the escapee, the pursuer using the improved algorithm for training can get closer to the escapee more quickly, and the average following distance also decreases.

## 1. Introduction

The research of multirobot system has been widely applied in industrial manufacturing, medical health, military operations, and other fields. The multirobot system [1] first appeared in the 1970s, due to the demand of practical application is increasing, it is difficult for a single robot to handle a lot of complex tasks, researchers have gradually turned their attention to multirobot systems. Unlike single robot system, multiple robots can share information to learn better strategy to accomplish cooperative tasks. The learning mechanism in the multirobot system enables robots to acquire environmental perception, information memory, and behavioral decision-making abilities in the process of interacting with the environment.

The multirobot systems application is pursuit-evasion problem that is a typical problem in the study of multiagent cooperation and coordination strategies [2]. In essence, the pursuit-evasion problem is a multiagent cooperative

decision problem. Robots study the optimal decision algorithm in multirobot system through the cooperative competition between the pursuers and the evaders. The methods to solve the problem of multirobot pursuit-evasion are mainly divided into the difference method and the combination method. The former converts the physical constraints of the robot into differential constraints and models and can only obtain local optimal solutions, but this method is widely used in solving the “simultaneous arrival” problem. [3]. The idea of latter method is to convert the pursuit problem into graph theory or dynamic programming problem and use greedy optimal return [4, 5], model predictive control [6], and other methods to learn the optimal strategy of the robot. Greedy algorithm [7, 8] is a typical combination method to fulfil the distance assignment task. This algorithm mainly realizes the allocation by minimizing the distance between the robot and the target. The application of this method is simple and extensive, but there will be a situation of allocation “deadlock.” Reference [9] uses the



negotiation method [10] to redistribute the coincidentally assigned round-up points that effectively settle the problem of deadlock. Traditional methods usually require complex mathematical calculation process to solve the multirobot pursuit-evasion problem. With the development of artificial intelligence, more researchers adopt machine learning methods to seek optimal chasing strategies.

Deep reinforcement learning (DRL) is a new machine learning method combining deep learning (DL) and reinforcement learning (RL). Compared with supervised and unsupervised learning, deep reinforcement learning combines the perceptual ability of deep learning with the decision-making ability of reinforcement learning [11] providing a solution to the complex and uncertain environment problems in multirobot systems [12, 13]. Volodymyr et al. first proposed deep Q-learning network (DQN) and obtained superhuman performance on Atari video games [14]. DQN is only applicable to discrete action space, and deep deterministic policy gradient (DDPG) is proposed to settle continuous action space issues [15]. Multiagent DDPG (MADDPG) is a multiagent policy gradient algorithm where agents learn a centralized critic based on the observation and actions of all agents [16, 17]. This method has already applied in the field of multirobot system. Kwak et al. [18] used reinforcement learning to train multirobot systems to obtain the optimal pursuit time. Li et al. [19] combined the data mining method of association rules with reinforcement learning and proposed a segmented reinforcement learning method, which solved the chase problem in the known environment. Liu et al. [20] used layered reinforcement learning to train the pursuit robot and designed an independent learning mechanism for the upper and lower stages [20] to balance the value function of different stages. The deep reinforcement learning method formulates an end-to-end strategy and is more targeted than traditional methods. However, the problem of inconsistent pursuit strategies in the phased training still exist. Additionally, the rewards of multirobot hunting scenes are sparse, and the algorithm stability performance is not satisfactory.

This paper considers the pursuit-evasion problem and uses a deep reinforcement learning technique to solve this problem. The sparse rewards and excessive exploration affect the training process of the robot, resulting in the failure to complete the pursuit mission. This paper proposes an improved MADDPG multiagent decision-making algorithm via combining environmental rewards and intrinsic rewards. The purpose of this paper is to solve the problem of excessive exploration after adding intrinsic rewards and sparse environmental rewards in multirobot pursuit-evasion. This method adds a curiosity module based on state similarity to the robot's reward function, under the premise of ensuring that the pursuit task is completed, a greater cumulative return is obtained. In addition, a threshold for similarity constrains the amount of intrinsic reward to overcome the problem of overexploration. The idea of this method is to combine environmental rewards and intrinsic rewards to facilitate the robot to learn better policies. This idea is maybe a general idea for improving multiagent decision-making algorithms.

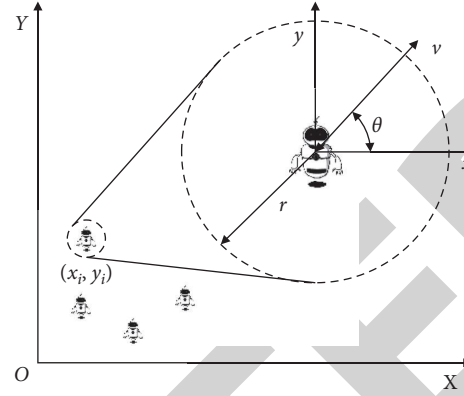


FIGURE 1: Robot motion model.

## 2. Multirobot Chase Model

**2.1. Robot Motion Model.** In the multirobot pursuit-evasion problem, this paper treats the individual robot as a particle, and the motion model of the robot is established in a two-dimensional plane. Both pursuers and evaders are randomly distributed in a 2D plane, and the dynamics model is the same for each robot. The information for each robot includes coordinate, velocity, direction, and detection radius, as shown in Figure 1. The kinematics model of the  $i$ th robot is expressed by where  $(x_i, y_i)$  represents the coordinate position of the  $i$ th robot,  $r_i$  is the radius of the sensing field of the  $i$ th robot,  $\theta_i$  is the included angle between the heading and the horizontal direction of the robot,  $v_i$  represents the robot speed along the heading,  $u_i$  is the change rate of the heading angle of the  $i$ th robot, and  $f_0$  is the damping coefficient.

$$\begin{cases} x = x_i, \\ y = y_i, \\ r = r_i, \\ v_{x_i} = v_i \cos \theta_i, \\ v_{y_i} = v_i \sin \theta_i, \\ \dot{\theta} = u_i, \\ f = f_0. \end{cases} \quad (1)$$

**2.2. Multirobot Cooperative Pursuit Problem.** Multirobot cooperative pursuit is to study how to instruct a group of autonomous mobile robots (pursuers) to cooperate with each other to chase another group of mobile robots (evaders), which involves cooperation between pursers and confrontation between pursers and evaders [21]. Compared with single robot pursuit, the uncertainty of multirobot system environment and the diversity of information bring great difficulties to the pursuit task. The multirobot pursuit scenario is shown in Figure 2. The red circle represents the pursuer, the blue circle represents the evaders, and the gray area represents the obstacle. A collision occurs when the distance between the center of mass of the two objects is less than the sum of the radii of the two objects. The red robot

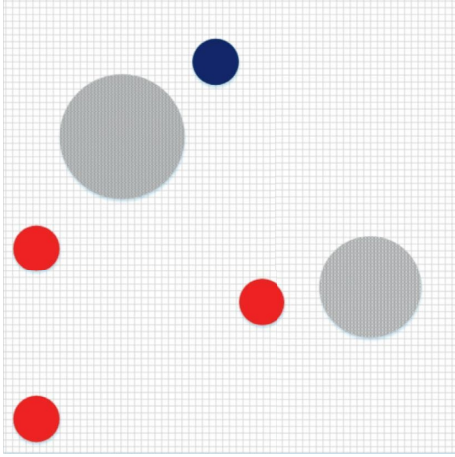


FIGURE 2: Schematic diagram of multirobot chase and escape scene.

needs to get as close as possible to the blue robot without collision between red robots to accomplish the pursuit task. The distance between red robot and blue robot is an indicator to measure the quality of the strategy. Red robot collides with blue robot is the best case but most case is that red robot follows blue robot keeping a certain distance.

There are some constraints in the multirobot pursuit-evasion scene. According to the robot motion model and multirobot pursuit process, the constraints are divided into control quantity constraint, initial state constraint, and distance constraint. Among them, the distance constraint including the obstacle constraint in the pursuit process of the robot and the safety distance constraint between the robots. The obstacle constraint requires no collision between robot and obstacle. Likewise, the safety distance constraint requires no collision between robots.

The control quantity constraint is reflected in the constraint on the change rate of the heading of each robot  $u_i$ .  $u_{i,\max}$  is the change rate of the maximum heading angle allowed by each robot, and the control quantity constraint of the  $i$ th robot is described as

$$|u_i| \leq u_{i,\max}. \quad (2)$$

The initial state includes the initial position and the initial heading angle, which are  $(x_0, y_0)$  and  $\theta_0$ , respectively. The initial constraint state corresponding to the  $i$ th robot is shown as

$$\begin{aligned} x_i(t_1) &= x_0, \\ y_i(t_1) &= y_0, \\ \theta_i(t_1) &= \theta_0, \end{aligned} \quad (3)$$

where  $x_i(t_1)$ ,  $y_i(t_1)$ , and  $\theta_i(t_1)$  represent the coordinates and heading angle of the  $i$ th robot at the starting time  $t_1$ .

The distance constraint is embodied in the distance between robots and the distance between robots and obstacles. There are  $N$  obstacles as shown in the gray shadow part in Figure 2. In the process of cooperative pursuit, any pursuer must not collide with obstacles at any time and maintain a certain safe distance. Assuming that

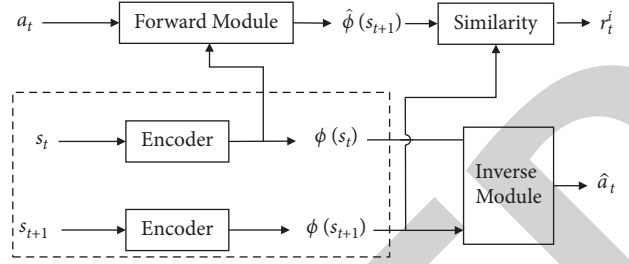


FIGURE 3: Curiosity module based on state similarity.

$D_i = [x_i, y_i]^T$ , the obstacle constraint of the  $i$ th robot is expressed as

$$\begin{aligned} \|D_i - D_n\|_2 &\geq r_n, \\ \forall i \in \{1, 2, \dots, S\}, \forall n \in \{1, 2, \dots, N\}, \end{aligned} \quad (4)$$

where  $n$  is the number of the obstacle;  $S$  is the total number of pursuers;  $N$  is the total number of obstacles;  $D_n$  is the center coordinate of the  $n$ -th obstacle; and  $r_n$  is the maximum radius of the  $n$ th obstacle.

The safe distance of each pursuer is considered to avoid collision occurs. Pursuers need to meet the following constraints between each other described as

$$\begin{aligned} \|D_i(t) - D_j(t)\|_2 &\geq R_s, \\ \forall t \in [t_1, t_f], \forall i, j \in \{1, 2, \dots, S\}, i < j, \end{aligned} \quad (5)$$

where  $R_s$  is the safe distance between pursuers and  $D_i(t)$  and  $D_j(t)$  are the coordinates of  $i$ th and  $j$ th robot at time  $t$ .

### 3. Methods

**3.1. Curiosity Module Based on State Similarity.** Intrinsic curiosity module (ICM) [22] is a solution to the sparse reward and insufficient exploration motivation in multi-agent environment. The error between the real state and the predicted state at the next moment is regarded as the intrinsic reward of agents in ICM to encourage agents to explore. According to the principle of intrinsic rewards, the larger the error between predicted state and real state, the more rewards agents obtain. Obeying this rule, supposing that the agent is exploring against real state direction at some point, it will always in wrong direction driven by abundant intrinsic reward that leads failure of pursuit task. Moreover, insufficient or excessive curiosity reward leads to unsatisfactory results [23]. Therefore, this paper adds a curiosity reward mechanism based on state similarity with threshold constraint on the basis of curiosity module. The degree of similarity between the predicted state and the real state represents the size of the prediction error, and the amount of intrinsic rewards to agents is effectively controlled by the constraint interval of similarity. When the similarity between the predicted state and the real state is below the lower threshold, agents cannot get internal excitation in order to prevent overexploration. In analogous cases, similarity exceeding higher threshold results in ignorance of intrinsic reward.

The intrinsic curiosity module is composed of three parts of neural network, encoder module, forward state prediction network, and reverse action prediction network, as shown in Figure 3. The encoder is used for agent state feature extraction, and the forward module makes use of the current state  $\mathcal{O}(s_t)$  and the current action  $a_t$  to estimate the state of the next generation  $\widehat{\mathcal{O}}(s_{t+1})$ . The inverse module employs a self-supervised style of predicting the current action by projecting forward the two  $\mathcal{O}(s_t)$  and  $\mathcal{O}(s_{t+1})$  outputs of the network to predict action  $\widehat{a}_t$ . The output of ICM is the two norm of  $\widehat{\mathcal{O}}(s_{t+1})$  and  $\mathcal{O}(s_{t+1})$ . Intrinsic rewards in this way do not take into account the problem of overexploration. Therefore, this method adds a similarity module to overcome this issue, as shown as Figure 3.

The state similarity module measures the degree of similarity between  $\widehat{\mathcal{O}}(s_{t+1})$  and  $\mathcal{O}(s_{t+1})$  by the Pearson coefficient (as shown in equation (6)). At the same time, this method overcomes the under-reward and over-reward by the threshold constraint of similarity. The intrinsic rewards produced by state similarity module is described as equation (7), where  $P$  is the Pearson correlation coefficient between  $\widehat{\mathcal{O}}(s_{t+1})$  and  $\mathcal{O}(s_{t+1})$  and  $\mu$  is the reward coefficient. Agents only gain intrinsic reward when Pearson coefficient is between *low* and *high*, *low* is the lower boundary of  $P$ , while *high* is the upper boundary of  $P$ .

$$P(x, y) = \frac{E[(x_i - \bar{x})(y_i - \bar{y})]}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (6)$$

$$r_t^i = P * \mu \quad (\text{low} < P < \text{high}). \quad (7)$$

**3.2. Improved MADDPG.** Multiagent reinforcement learning algorithm is an effective method to settle the problem of multirobot pursuit-evasion. Each robot's decision is controlled by its own neural network, and the pursuers can complete the task of chasing fugitives through communication interaction and cooperative decision. MADDPG algorithm is improved on the basis of DDPG algorithm by Lowe ([16]) and applied to multiagent decision-making scenarios. Its core idea is to find the optimal global strategy by means of centralized training and distributed execution, which solves the problems of nonstationarity and the failure of experience replay method in multiagent environment. However, robots lack exploring motivation due to sparse rewards, which causes low rewards and model difficult to converge.

This paper combines intrinsic reward and external environment reward as the agents' overall rewards via introducing the curiosity module based on state similarity into MADDPG. In this way, this method not only simulates the enthusiasm of the robot to study strategy but also prevents the excessive behavior. The model convergence speed and cumulative return value of the model are significantly improved using this method. Multiagent training is a process of searching for strategies that generate maximum total reward expectation and is described as shown in equation (8).

$$\max_{\theta_p} E_{\pi(s_t; \theta_p)} \left[ \sum_t r_t \right], \quad (8)$$

where  $\theta_p$  is the neural network parameter and  $r_t$  is the total reward value of the agent at time  $T$ . The total reward value consists of two parts, namely, the internal motivation is  $r_t^i$  and the external environment motivation is  $r_t^e$ , as shown in equation (9). In general, the intrinsic motivation is 0. When the intrinsic curiosity module is added, the intelligence will take the similarity error between the predicted value and the actual value of the next state  $s_{t+1}$  as the intrinsic reward according to the current state  $s_t$  and the given action  $a_t$ , as shown in

$$r_t = r_t^i + r_t^e. \quad (9)$$

Improved MADDPG algorithm training framework adds an internal curiosity module on the basis of MADDPG, and its overall architecture is shown in Figure 4. As a whole, the actor-critic framework is still adopted [24, 25] Each agent is training by Actor network, Critic network, Target Actor network, and Target Critic network. The parameters of actor network and critic network are updated via gradient descent computing the overall loss. The parameters of the target network update via copying the parameters of the local network through soft update within a certain period of time.

This method adds a module named ICM, as shown in Figure 4. The inputs to this module are current real state  $S_t$ , next real state  $S_{t+1}$ , and current action  $a_t$  of each robot. The output is the intrinsic reward generating by state similarity module. The sum of intrinsic reward  $r_t^i$  and external reward  $r_t^e$  is put into replay buffer to update network. Through the optimized reward calculation method, the improved algorithm gives appropriate rewards to the robot resulting in better strategy that guides robots to catch up on the target robot faster and more accurately.

**3.3. Algorithm Process.** The improved algorithm flow is shown in Table 1. Before training of agents, the replay buffer and exploration noise have been initialized, the network parameters of actor, critic, and the additional curiosity module also have been initialized. Every robot's state is reset at the begin of each episode, a quadruple  $s_t, a_t, r_t, s'_t$  will update after each robot chooses an action and then is put into replay buffer,  $s_t$  and  $s'_t$  are robot's current state and next state,  $a_t$  is current action of robot, and  $r_t$  is the whole reward of robot. The value network and target network parameters will be updated by sampling in the replay buffer.

## 4. Design of the Multirobot Pursuit Strategy

**4.1. State Space.** For the multirobot pursuit-evasion scene, both the pursuer and the evaders are treated as particles, and the state space of the pursuit scene is divided into the pursuer and the evader.

The pursuer's state space includes the pursuer's centroid position  $c(t)$ , the distance  $d(t)$  from other pursuer, the distance  $l(t)$  from the obstacle's centroid, the distance  $w(t)$

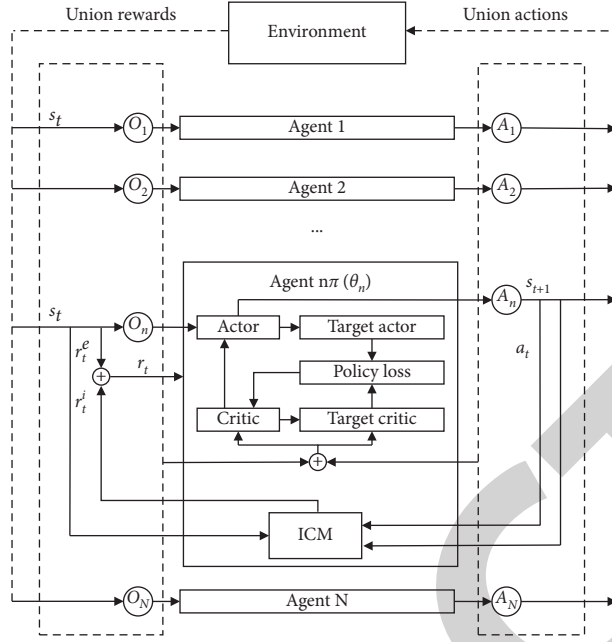


FIGURE 4: Improved MADDPG algorithm framework.

TABLE 1: The multirobot collaborative pursuit strategy using improved-MADDPG.

- (1). Initialize replay buffer  $D$  and motion exploration noise  $\varepsilon$
- (2). Initialize the EvalActor network and EvalCritic policy network for each robot
- (3). Initialize the TargetActor network and TargetCritic network for each robot
- (4). Initialize the ICM network for each robot
- (5). For episode = 1 to MaxEpisode:
  - (6). Initialize the status of chase robot and escape robot  $s_t$
  - (7). For  $t = 1$  to MaxStep:
    - (8). Each chaser and runner chooses the action  $a_t$
    - (9). Obtain  $s_t, a_t, r_t^e, s_t'$
    - (10). Calculate the intrinsic reward  $r_t^i$  by  $s_t$  and  $s_t'$  similarity error  $P$
    - (11). Update overall rewards  $r_t = r_t^i + r_t^e$
    - (12). Update current status  $s_t'$ , push  $(s_t, a_t, r_t, s_t')$  into replay buffer  $D$
    - (13). For agent = 1 to  $N$ :
      - (14). Take  $s$  samples from the experience pool to form batch samples, and calculate union losses to update EvalCritic network parameter
      - (15). Soft update the TargetCritic network parameter
      - (16). End for
      - (17). Soft update TargetActor and TargetCritic network parameter
      - (18). End for
      - (19). End for

from the evader, the heading angle  $\varphi(t)$ , and the heading speed  $v(t)$ . The pursuer's state space at time  $t$  is defined as follows:

$$s_p(t) = [c(t), d(t), l(t), w(t), \varphi(t), v(t)] = \begin{bmatrix} (x, y), (d_1, d_2, \dots, d_N), (l_1, l_2, \dots, l_M), (w_1, \dots, w_K), \\ \varphi, (v_x, v_y) \end{bmatrix}, \quad (10)$$

where  $(x, y)$  is the pursuer's coordinate,  $(d_1, d_2, \dots, d_N)$  is the distance between the current pursuer and peer pursuers,  $N$  is the number of peer pursuers,  $(l_1, l_2, \dots, l_M)$  is the

distance between the pursuer and the center of mass of the obstacle,  $M$  is the number of obstacles,  $(w_1, \dots, w_K)$  is the distance between the pursuer and the evader,  $K$  is the

number of evaders,  $\varphi$  is the heading angle, and  $v_x, v_y$  are the velocities along the  $x$  and  $y$  directions, respectively.

The state space of the evader includes the position  $c(t)$  of the center of mass of the evader, the distance  $d(t)$  between

$$s_b(t) = [c(t), d(t), l(t), w(t), \varphi(t), v(t)] = \left[ \begin{array}{c} (x, y), (d_1, d_2, \dots, d_{K-1}), (l_1, l_2, \dots, l_M), (w_1, \dots, w_{N+1}), \\ \varphi, (v_x, v_y) \end{array} \right]. \quad (11)$$

**4.2. Action Space.** This paper regards multirobot pursuit-evasion as a continuous process. Therefore, the action space is complex and continuous rather than simple and discrete. In accordance with the movement characteristics of the robot, the action space at time  $t$  is designed, as shown in

$$\begin{aligned} A(t) &= \{\Delta\varphi, \Delta v\} \\ &= \{\varphi(t) + \mu t, v(t) + at\}, \end{aligned} \quad (12)$$

where  $\Delta\varphi$  is the heading angle change value,  $\mu$  is the angle change rate,  $\Delta v$  is the heading velocity change value, and  $a$  is the acceleration. The movement process of the robot is mainly related to the speed and direction, angle change rate controls direction, and acceleration controls velocity.

**4.3. Reward Function.** With the design idea of improved MADDPG algorithm, the reward function is divided into two parts, including external environment reward and internal incentive reward.

In the multirobot pursuit-evasion scenario, pursuer is rewarded for catching up evaders. On the contrary, evaders who are caught are punished. However, sparse reward is not conducive to improving the comprehensive performance of the pursuer. Therefore, distance reward is also taken as an external reward in the design of reward function in this paper.

The relative position reward of the pursuer and the evader can be described, as shown in equation (13),  $\omega$  is the relative distance reward coefficient,  $M$  is the number of pursuers,  $N$  is the number of evaders,  $x_{\text{predator}}, y_{\text{predator}}, x_{\text{prey}}, y_{\text{prey}}$  are the  $x$  and  $y$  coordinates of the pursuer and the evader, respectively.

$$r_t^e = \omega * \sum_{i=1}^M \sum_{j=1}^N \left\| (x_{\text{predator}}^i, y_{\text{prey}}^j) - (x_{\text{predator}}^j, y_{\text{prey}}^i) \right\|. \quad (13)$$

The reward of the pursuer for completing the pursuit task is described, as shown in equation (13).  $\delta$  is the reward base value and the pursuer and the evader get negative reward.

$$r_t^e = \begin{cases} +\delta, & \text{if agent is predator,} \\ -\delta, & \text{if agent is prey.} \end{cases} \quad (14)$$

As for intrinsic reward, it is designed to encourage robots to learn optimal strategy. This paper describes intrinsic reward as equation (15),  $\mu$  is the reward coefficient and the

the evader and fellow evaders, the distance  $l(t)$  between the center of mass of the obstacle, the distance  $w(t)$  between the pursuer, the heading angle  $\varphi(t)$ , and the heading speed  $v(t)$ . The state space of the evader at time  $t$  is defined as follows:

intrinsic reward is generated by similarity error between the predicted state  $\hat{s}_{t+1}$ , and current state  $s_{t+1}$  at time  $t + 1$ .

$$r_t^i = \mu * P(\hat{s}_{t+1}, s_{t+1}). \quad (15)$$

The total reward  $r_t$  is the sum of external environment reward and internal incentive reward described as

$$r_t = r_t^e + r_t^i. \quad (16)$$

## 5. Simulation Experiment

**5.1. Experimental Configuration.** This paper adopts the hypothetical scenario in Figure 2 to complete a series of experiments, in which the pursuers cooperate to chase the evaders. The starting point of the task is when the pursuers start to move, and the end point is when they catch up with all the evaders. The simulation environment was redesigned on the basis of MPE environment and Python. OpenAI Gym reinforcement learning platform was adopted, and Pytorch deep learning framework was used to complete the simulation experiment of multirobot cooperative pursuit task.

The training of robots is a close-loop process, robots obtain environmental information by taking actions through strategy they have learned to update their own status information, then optimize pursuit-evasion strategy by means of the feedback information. The hyperparameters in the experiment are shown in Table 2. The whole training process iterates  $2 * 10^5$  times, each episode includes 100 steps. The critic network learning rate is slightly higher than actor network learning rate so that critic correctly guides actor. Discount factor represents the degree to which each training session is learned from the previous experience. The size of replay buffer and batch size depend speed of experience feedback.  $\Delta T$  is the time interval for each simulation playback.

Different from deep learning, deep reinforcement learning has no data set. The agent obtains data from the simulation environment and uses the deep neural network for strategy optimization to train the model. Combined with the characteristics of deep reinforcement learning and the actual multirobot pursuit-evasion scenario, the following assumptions are made for the simulation environment:

- (1) To prevent the motivation generated by the internal curiosity module from keeping the robot exploring without stopping, this method limits environment boundaries and maximum stride length per turn.



TABLE 2: The training hyperparameters.

Training parameter	Values
Max_steps	$2 * 10^5$
Maximum episode length step_len	100
Discount factor $\gamma$	0.96
Actor network learning rate $\alpha_a$	$10^{-4}$
Critic network learning rate $\alpha_c$	$10^{-3}$
Size of replay buffer $M$	$10^6$
Batch size	256
Simulation time step $\Delta T$	0.1 s

- (2) All pursuers have exactly the same parameters, and they communicate with each other by share status information.
- (3) Considering the characteristics of deep reinforcement learning, the robot's range of activity is a two-dimensional infinite space. But the robot only moves within a certain range in practice, thereby this paper limits the multirobot pursuit scene to a two-dimensional finite space in that the robot is randomly initialized with coordinates between  $-1$  and  $1$ . Robots who are beyond the limited range are given certain punishment, and the environment is composed of  $N$  pursuers,  $K$  evaders, and  $M$  obstacles.

**5.2. Experimental Analysis.** Deep reinforcement learning models usually judge algorithm performance from two aspects: convergence speed on the one hand and reward value after convergence on the other hand. In the experiment, the algorithm performance of MADDPG and the improved MADDPG algorithm in the multirobot pursuit-evasion problem are compared. Those methods including MADDPG, MADDPG with ICM, and improved MADDPG are tested in the 3VS1 scenario. The curve between the robot's average reward value and the number of iterations is drawn to evaluate the performance of algorithms, as shown in Figure 5. The reward value of the robot is the most intuitive manifestation of the performance of the algorithm. In the multirobot reinforcement learning training process, the average return value of the robot is calculated after each training episode to measure the quality of strategy and whether the algorithm has converged is judged by observing the training curve.

Figure 5 indicates that the overall average reward of the improved algorithm is significantly improved, and the method of adding state predicts error into intrinsic curiosity module accelerates the model convergence speed and has better convergence stability.

From the trend of the curves in Figure 5, the convergence value of the black curve (i.e., the MADDPG algorithm) is significantly lower than that of the other two curves. The red curve shows the trend of the reward value after adding the original curiosity module. The reward is increasing compared with MADDPG under the influence of increasing explore motivation via intrinsic reward generation by ICM. However, in the convergence stage, the curve oscillates greatly caused by excessive exploration. When the rules of

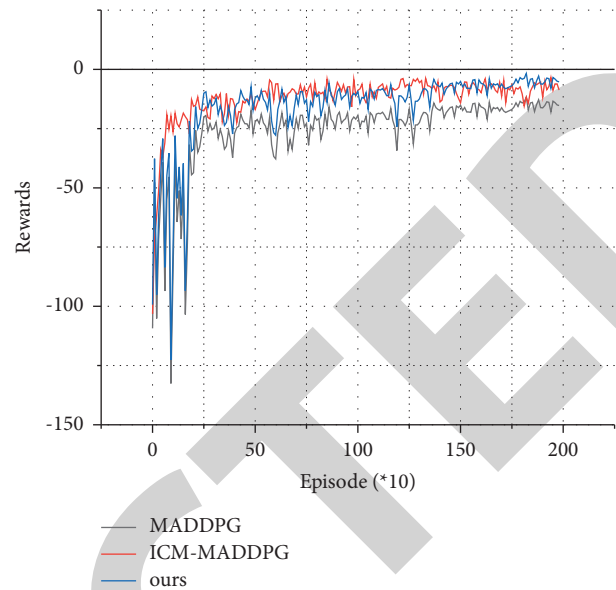


FIGURE 5: Average reward value of simulation scenarios.

giving intrinsic rewards are changed, the threshold of state similarity module limits intrinsic reward produced by ICM and guides the robot to explore rationally. Although the overall reward value is not greatly improved, the stability in the convergence stage is improved as shown by the blue curve.

In addition, in order to test the stability of the algorithm model, the trained model was tested in the 3VS1 scenario, and the average reward and task completion rate under 1000 tasks were counted. The results are shown in Figure 6 and Table 3.

The test results shown in Figure 6 show that the improved MADDPG algorithm has significantly improved in both average reward value and task success rate. Taking task completion rate as a performance indicator, the improved algorithm has improved the success rate by 2–4 percentage points compared with the original algorithm.

Figure 6 shows the specific experimental data of the three algorithms in the form of a bar chart. Figure 6(a) shows the average reward of the robot, and Figure 6(b) shows the success rate of the multirobot system under the three algorithms. Table 3 indicates that both reward and success rate are increasing and proves the effectiveness of the improved algorithm.

In the pursuit-evasion scenario of 3VS1, the distance between the pursuer and the evader is also an important indicator to measure the performance of the algorithm. Figure 7 shows the track graph of a pursuit-evasion process and the real-time distance trend graph of the pursuer and the evader. The improved MADDPG algorithm enables the robot to follow the target more quickly and get closer to the target on average within the specified time. The statistical results are shown in Table 4.

The four pictures in Figure 7 intuitively show the robot's pursuit-evasion process, where Figures 7(a) and 7(c) are the trajectories of the robot tested using MADDPG and the



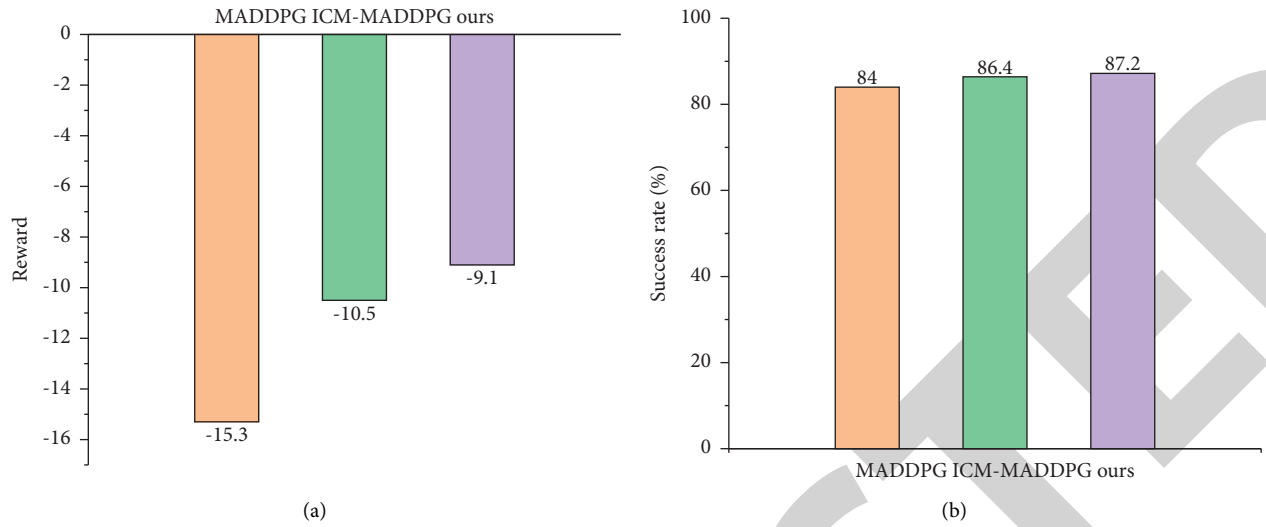


FIGURE 6: 1000-times test results. (a) Average reward. (b) Success rate.

TABLE 3: Average reward value and success rate.

Algorithm	Average	Success rate (%)	Improvement (%)
MDDPG	-15.3	84	—
ICM-MDDPG	-10.5	86.4	2.85
Ours	-9.1	87.2	3.81

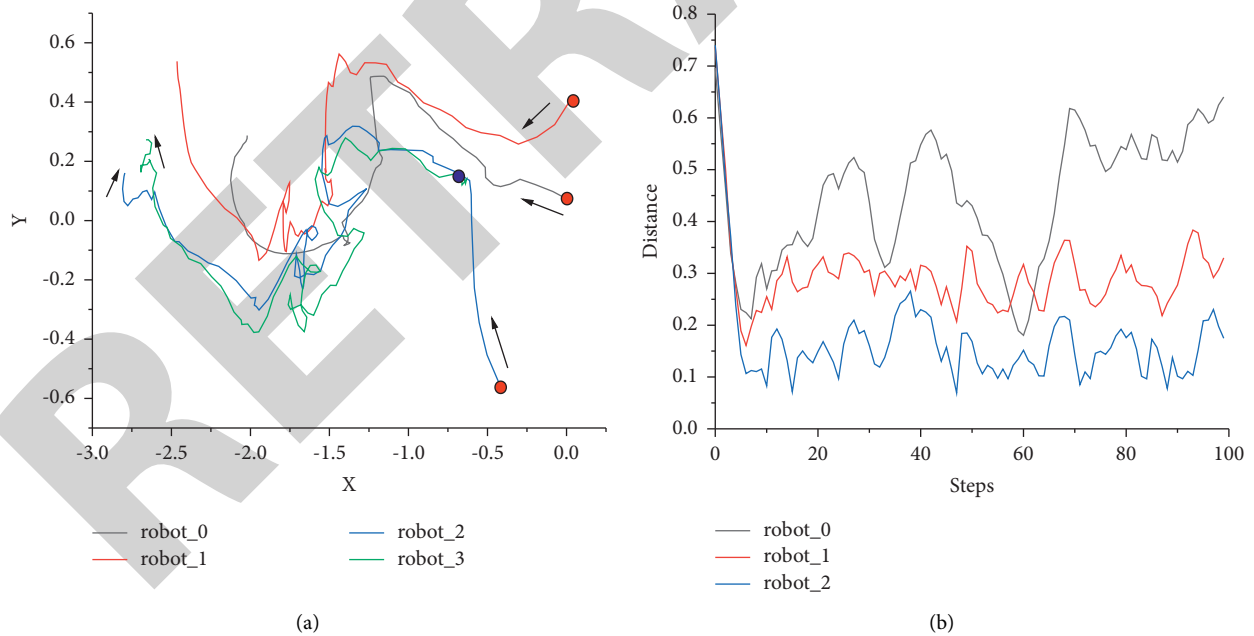


FIGURE 7: Continued.

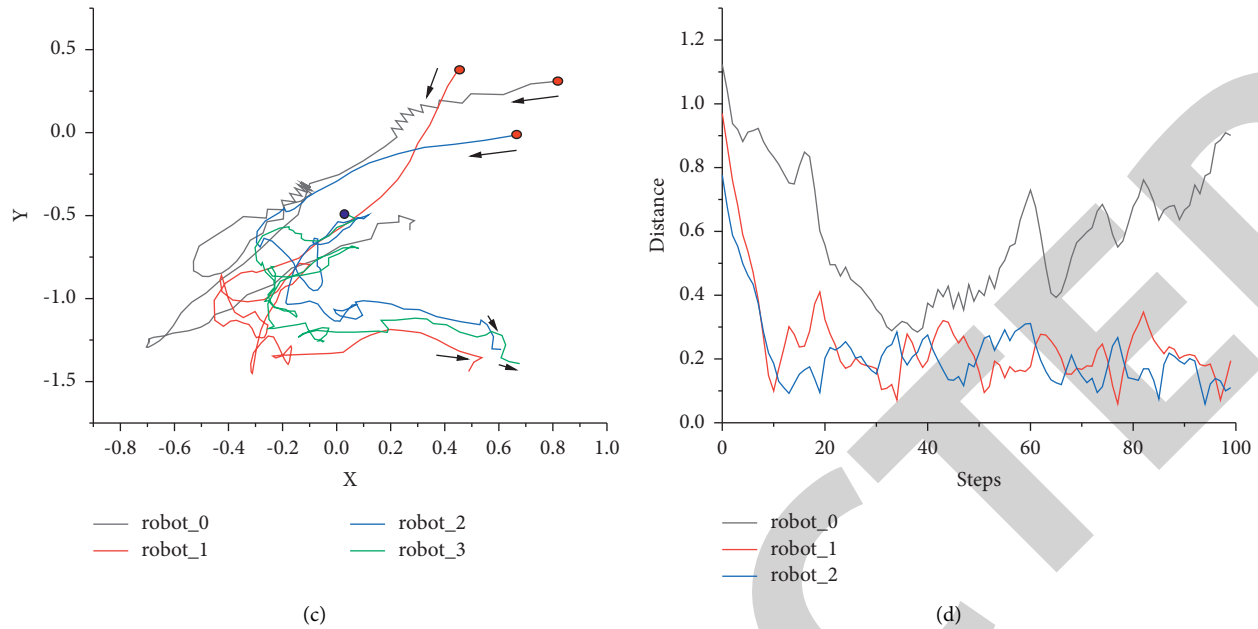


FIGURE 7: Robot trajectory and real-time distance in the 3VS1 scenario. (a) Trajectory of the robot using MADDPG; (b) real-time distance between the pursuer and escaper using MADDPG; (c) trajectory of the robot using improved MA-DDPG; (d) real-time distance between the pursuer and esca-per using improved MADDPG.

TABLE 4: Average distance between the pursuer and evader in the 3VS1 scenario.

Algorithm	Average distance		
	robot_0	robot_1	robot_2
MADDPG	0.56	0.30	0.15
Ours	0.73	0.21	0.14

improved algorithm. It is obvious that the improved algorithm drives more robots to approach the target robot, and their ability to follow the target is significantly improved. Figures 7(b) and 7(d) describe the real-time distance between the pursuer and the evader during the entire pursuit-evasion process. The average distance between the pursuer and the evader in a test is shown in Table 4. The improved algorithm can reduce the average distance between robot\_1 and robot\_2 and the target robot, which represents an improvement in the performance of robot hunting. Although robot\_0 performs poorly, it does not affect the success of this pursuit task. One possible reason is that the robot sacrifices itself to provide other partners with richer information, thereby making other robots faster approach target.

## 6. Conclusion

To address the problem of changeable environment and sparse rewards in the process of multirobot pursuit-evasion, this paper proposes an intelligent multirobot pursuit-evasion method based on deep reinforcement learning. A multirobot pursuit-evasion model was built to describe motion model of robot and some constraints in the multirobot pursuit-evasion scene. The improved

curiosity module with state similarity based on threshold constrain is introduced into the MADDPG algorithm to resolve sparse reward and unstable model caused by excessive exploration. The improved MADDPG algorithm is better applied to the multirobot pursuit-evasion scene by designing the state space, action space, and reward function of the agent. Comparing the improved MADDPG algorithm with the MADDPG algorithm, the results show that the improved algorithm can not only solve the problem of multirobot pursuit-evasion but also the average reward of the agent is obviously higher than the original algorithm.

At present, the improved algorithm has better results for specific scenarios, the next step is to further optimize the algorithm so that it can play a better role in different scenarios.

## Data Availability

No data set was used in this paper, and the data used in reinforcement learning were obtained by the agent from the simulation environment.

## Conflicts of Interest

The authors declare that there are no conflicts of interest.

## Acknowledgments

This work was sponsored by the National Key Research and Development Project (2021YFC3001502) and the National Natural Science Foundation of China (52072292).

## *Retraction*

# **Retracted: Application of Mathematical Models in Economic Variable Input and Output Models under the Scientific Visualization**

### **Computational Intelligence and Neuroscience**

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### **References**

- [1] K. Zheng and Y. Liu, "Application of Mathematical Models in Economic Variable Input and Output Models under the Scientific Visualization," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 6269358, 10 pages, 2022.

## Research Article

# Application of Mathematical Models in Economic Variable Input and Output Models under the Scientific Visualization

Kai Zheng<sup>1</sup> and Yisheng Liu <sup>2,3</sup>

<sup>1</sup>Research Institute for the Development of Private Economy, A New Type of Think Tank with Characteristics in Colleges and Universities in Fujian Province, Fujian, Quanzhou, China

<sup>2</sup>College of International Finance and Trade, Zhejiang Yuexiu University of Foreign Languages, Shaoxing, Zhejiang, China

<sup>3</sup>School of Political and Social Development, Quanzhou Normal University, Quanzhou, Fujian, China

Correspondence should be addressed to Yisheng Liu; liuyisheng200208@sina.com

Received 11 January 2022; Accepted 26 January 2022; Published 21 February 2022

Academic Editor: Daqing Gong

Copyright © 2022 Kai Zheng and Yisheng Liu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

At present, with the development of society and economy, some new problems have emerged continuously. Among them, the more serious problem is that enterprises pay too much attention to economic benefits, which leads to problems in the development of many enterprises. Therefore, the problem caused by too much emphasis on economic benefits is one of the major economic problems. After this problem, our country's economy began to focus on economic input and output in the subsequent development process. And, related personnel have also studied the economic input-output technology and model, which is an economic quantitative analysis method that has been widely used in various aspects of the economic field since its emergence. This paper firstly elaborates the theoretical knowledge involved, such as economic data, economic variables, and visualization techniques. Secondly it analyzes the application of scientific visualization techniques in the processing of economic variable data, which includes the techniques of preprocessing, mapping, drawing, and displaying the data. Finally, this paper investigates the application of mathematical models in economic variables input and output and the application of mathematical models in economic research, which can help us to better understand and use the knowledge of economics, thus providing strong support for solving practical problems. Also, this helps people to understand and grasp more about the macroarea, microaspect, and even the country as a whole.

## 1. Introduction

Beginning in the late 1970s, some foreign researchers have made in-depth discussions on visualization-related issues, such as measurement, mathematical models, and so on. Since 1980, our country began to pay attention to the fields of scientific visualization theories and methods. In addition, many countries have done a lot of research on the economic input-output model, which makes the economic input-output model more widely used in the economic development of enterprises [1]. The input-output model is mainly used for the analysis of the chain impact of product price changes, the analysis of investment-induced effects, and the analysis of industrial structure adjustment. At present,

domestic scholars have not formed a systematic and effective framework to serve quantitative analysis. It is also an important problem to be solved at present to conduct a deeper research on how to establish the relationship between the input and output of a set of economic variables [2].

In summary, this paper mainly studies the following contents:

Mathematics is a highly abstract discipline that covers many fields and has a wide range of applications in economics, sociology, and psychology, among others. We can see that classifying variables from different perspectives often requires the consideration of several factors at the same time. For example, when people are faced with a problem in their lives, they choose to think about all the possible

solutions to the problem before making a decision [3]. Mathematical visualization is the use of computer technology to visualize areas that do not require precise data, such as functions and probabilities. It allows us to use existing knowledge to solve practical problems more easily, quickly, and effectively. The variables involved in this process can also produce a certain degree of “backward” effect or partially offsetting effect (which is unpredictability and nonordered error) with time change, thus affecting the effect of mathematical modeling, which makes the visualization theory more comprehensive and accurate to play its proper value [4, 5]. In this paper, we introduce some visualization variables based on previous research and then build a mathematical model and dig deeper into it.

Economic variables are important factors in explaining the production function and can generally be reflected by the ratio of a given output to income. In real life, there are usually one or more economic variables that are related to output, have some correlation, and can be expressed in different quantities. For example: when unit output value =  $GDP * (X1 + Y0)/y$ , the sales price of its product is zero; for the same linear equation, it is output per unit  $\times x$  value. That is, each production function has more than two given quantities and another known variable, and each output value has more than two known quantities and another unknown variable [6].

Mathematical models can concretize abstract, non-quantitative problems, and in practical applications it has a profound impact on people's lives. In the traditional method, some situations are unpredictable or difficult to explain by decision makers, and this method can be used to solve them. The word “precision” first appeared in the mid-19th century. With the progress of the times and the technological revolution, the word “precision” was introduced. A mathematical model is a tool for describing, reasoning, and calculating the relationships between things and the laws arising from changes in their nature by using certain quantitative and quantifiable objects. Precise mathematical models play an important role in economic research, which can help people to simplify complex problems and solve practical problems, so it is getting more and more attention and application [7, 8].

The application of mathematical models is mainly to solve the visualization of economic variables, and in reality, some data cannot be directly obtained, so it needs to be transformed into something that can visually reflect the essential characteristics and inner laws of things, reflect the nature of the problem, or describe things, and other functions are clearer and easier to understand and grasp the interconnection and influence relationship between the phenomena under study. This requires us to pay attention to the actual situation as well as to grasp the development trend of scientific knowledge itself. Mathematical models as an effective method can be used to solve some difficult problems in the visualization of economic variables, which are widely used in many fields [9].

Scientific visualization refers to the realization of human-computer interaction with the knowledge and information involved in the process of data acquisition,

processing, and analysis through computer technology, so that people can understand and use it. Essentially, it is a way of thinking. In traditional mathematical models, only the relationship between variables and the internal structure characteristics of variables are usually studied to describe the degree of influence of random events, which ignores its internal laws and state change trends. However, modern visualization can help researchers to understand and analyze the problems, causes, and development prospects in complex systems more deeply, so visualization technology is more and more widely used in mathematical research [10, 11].

The economic data is the basis for the construction of the visualization model. In the process of building it, it is necessary to classify the variables under study and then analyze the interrelationships and internal logical relations between these indicators. For example, for the attribute of production (output), cost of production, etc., it is a function, while for the attribute of consumer demand it is utility. These two sets of regression coefficients and test results are used to determine the level of economic development, how it determines the amount of input, and whether it grows and thus affects the rationality of the visualization modeling, which provide valid support for policy makers.

For mathematical visualization, scholars at home and abroad have carried out research from different angles and achieved some good results. These theories can be used to solve practical problems. There are many related discussions on science abroad. For example, two of the famous foreign economists Kahneman and Delphi put forward the “strategic” concept (Non-Levin); Professor James Gist of Harvard University in the United States believes that visualization is not only an information processing technology or tool, but also a source of human wisdom. The research on mathematical visualization in our country started late, and the research on visualization mainly focuses on economic variables, explanatory variables, and application fields. There is no complete system about how to apply the mathematical model to practice.

Scientific visualization refers to people's intuitive, concrete, and vivid definition of complex problems or phenomena and makes them a clear and easy-to-understand expression through visual and auditory methods. It mainly includes three aspects: the first is to make an accurate judgment on the existence of a large number of uncertain factors in the objective world; the second is to make an abstract generalization of the natural system in essence; the third is to describe the nature and various phenomena involved in the process of social development and their interrelations and interactions. The research model based on visual analysis is shown in Figure 1.

Based on the economic study of scientific visualization, we analyze the following aspects. Firstly, mathematical modeling: In the traditional field of economics, it is usually assumed that a variable can be expressed by a simple formula; secondly, the mathematical methods are improved and refined and applied to real life, converting the data that were impossible or partially unavailable to the results that can be obtained by simply calculating a functional

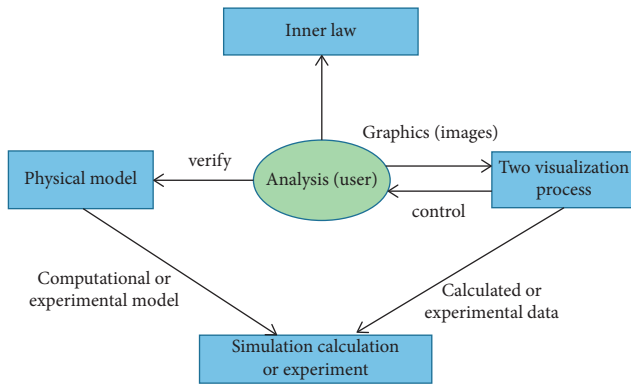


FIGURE 1: Research model based on visual analysis.

expression, which is called a visualization algorithm. Finally, computer-aided technology is used to combine economic indicators with other nonquantitative factors to obtain quantitative indicators [12].

## 2. Materials and Methods

**2.1. Data Preprocessing.** For data processing we usually use some mathematical methods, such as statistical models, image models, and so on. However, these common econometric research tools often require a large number of complex and detailed models of the sample in practice. Therefore, in this paper, we will select appropriate and reasonable information that can reflect the real situation and representative information as the essence, so as to better provide more useful information for decision makers to make accurate choices and reference data for decision making; at the same time, it can also provide more useful information for specific decision makers, so that they can make more accurate and rational choices. The block diagram of data preprocessing is shown in Figure 2.

In this paper, mathematical tools based on the “least squares” algorithm (SURF) and the LM heuristic algorithm are used to complete the data preprocessing process. Among them, the “least squares” algorithm (SURF) is mainly used for data correction; i.e., when there is only random error in the variables, as described in formula (1), and when there is significant error in the variables, the algorithm is described in formula (2).

$$x^m = x + \varepsilon. \quad (1)$$

$$x^m = x + \varepsilon + b. \quad (2)$$

One of the characteristics of display technology is the display of the amount of data that needs to be processed, especially image data. As the resolution of image capture tools increases, the amount of image data per frame increases exponentially. Therefore, the storage and transmission of data have a negative impact, which means that the capacity of internal/external memory must be expanded to increase the speed of data communication. One of the technical methods to overcome this difficulty is to

use data compression technology. Traditionally, data compression technology has always been one of the main research materials in digital image processing. Data compression technology can be divided into lossless compression and lossy compression according to the degree of distortion of information after compression. And its compression time is relatively long, but when the data block is processed and compressed, it is prone to overflow. The development of scientific computing screen technology and the development of multimedia computing technology have promoted the development of data compression technology. The current data compression technology can reduce the data storage capacity, but its computational complexity is high, and it needs continuous improvement. In recent years, data compression algorithms and standardization have developed rapidly, and integrated circuits and dedicated hardware used for compression applications have become the main topics of current research. The data compression technology flowchart is shown in Figure 3. The implementation of the data algorithm is in Algorithm 1.

Through this article, we can understand that data compression technology can effectively expand storage capacity and accelerate the speed of data communication; now there are various data compression systems on the market. The performance metrics should include (1) minimum error density; (2) maximum variance rate as well as minimum entropy value; (3) minimum generative model; 4–6 are the missing values of metrics; i.e., the weights and maximum entropy values are selected as the measurement factors without considering the nonordered visualization.

Data compression techniques can be divided into lossless compression and lossy compression according to the degree of distortion to the information after compression. Lossless compression technique is to segment data points with certain characteristics, which is characterized by reducing the amount of information required in random measurements; it is a probability density estimation technique; this can simplify the calculation of statistics. Lossy compression is characterized by a higher compression ratio than lossless compression, but it usually causes distortion after decompression. Also, the number of variables is uncertain and the variables are closely related to each other.

According to the output distribution characteristics of the compressed information source, data compression can be divided into statistical coding and dictionary coding. Statistical compression is one of the earliest compression techniques used and it is associated with classical data compression. Vocabulary encoding is also called LZ encoding. The basic principle is to select a character string from the source output data, save each character string code as an identifier, and use the dictionary search principle to encode the compressed character string and perform character string encoding.

Data compression technology can reduce data storage capacity. The principle of data compression technology is as follows: due to some redundancy between image data, this makes data compression possible. Shannon, the founder of information theory, proposed to treat data as a combination



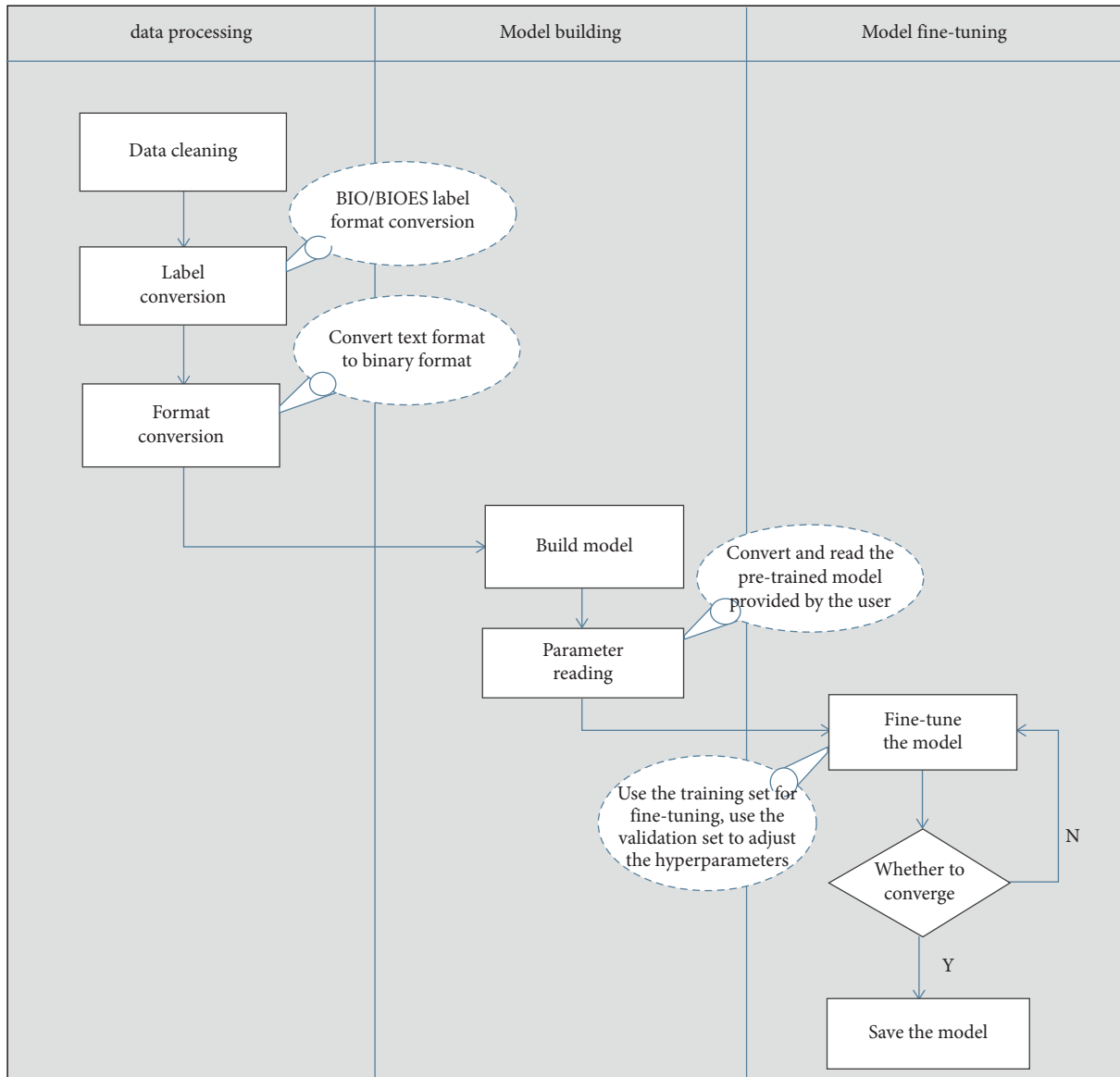


FIGURE 2: Data preprocessing flowchart.

of information and redundancy. The so-called redundancy is due to the high correlation between the pixels of the image, so many coding methods can be used to remove them, so there is no need to reduce the compressed data.

One of the characteristics of visualization technology is the display of the amount of data that needs to be processed. The application of visualization technology in mathematics is mainly reflected in, first, the integration of data and information, through the transformation of complex abstract problems into simple and clear, easy-to-understand-and-grasp features; second: the simplification of the model can be easily applied to many fields, for example, the relationship between variables in the function model and the correlation matrix representation between parameters. Third, by constructing

mathematical models, data information can be combined with the actual situation, thus realizing the characteristics of high accuracy of visualization.

**2.2. Mapping (Modeling).** The concept of mapping (PERT): it has the properties of fixed state, regular change, and uncertainty and satisfies the measurable characteristics; i.e., there is a certain number or correspondence between variables. When analyzing a mathematical model, it is important to understand the domain of definition and its properties [13]. For the problem we are studying, it can usually be seen as a constraint or a function to describe the existence of some connection between the state quantities; and when it comes to the actual situation that

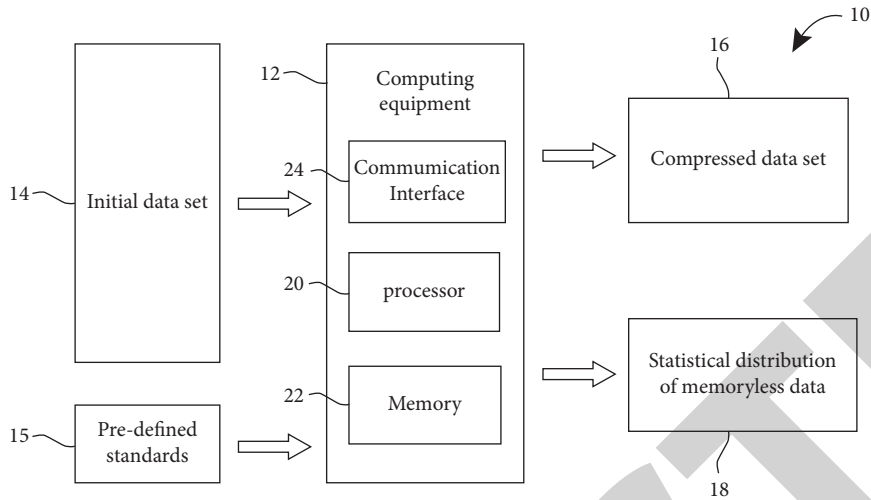


FIGURE 3: Flowchart of data compression technology.

```

int Rle_Encode_N (unsigned char * inbuf, int inSize, unsigned char * outbuf, int onuBufSize)
{
    unsigned char * src = inbuf;
    int i;
    int encSize = 0;
    while (src < (inbuf + inSize)){
        if((encSize+2) > onuBufSize)/*Not enough space in the output buffer*/
        {return -1;}
        unsigned char value = * src++; i = 1; while(src == value) && (i < 255)){
            src++;
            i++;
        }
        outbuf[encSize++] = i;
        outbuf[encSize++] = value;
    }
    return encSize;
}
    
```

ALGORITHM 1: Processing of economic variable data.

needs to calculate some specific values with definite integrals, the mapping (PERT) is used as one of the solutions, which is also an important prerequisite for the study of variables. For example, according to the mathematical function (3), the corresponding map can be drawn as shown in Figure 4.

$$y = \begin{cases} x^2 - 4x + 4, & (x \geq 2) \\ \frac{x}{2} - 1, & (x < 2) \end{cases} \quad (3)$$

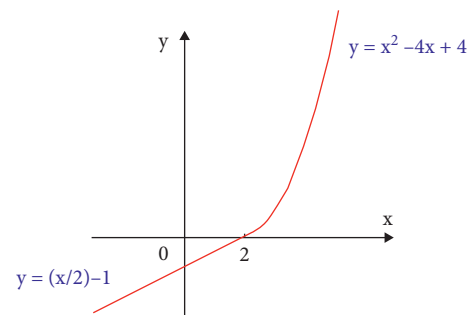


FIGURE 4: Function map.

2.3. *Drawing.* The drawing function can complete the process of converting any data into drawing data. As we all know, the principles and methods of adult computer graphics provide many rendering algorithms for display technology. Scan conversion includes hidden surface removal, light mode, shadow, transparency and shadow, texture mapping, and antialiasing technology. In general, the rendering method provided by the graphics computer can meet the needs of the rendering module in the display

technology, so the research on the rendering function is not the main issue of the display technology. However, as mentioned earlier, in some new fields of display research, rendering technology can also become an important research technology, such as volume rendering technology.

Volume rendering technology is a relatively important research technology, and volume rendering technology is based on the image lighting model. The basic

principle is a method based on optical diagrams. This method maps the three-dimensional data field into the basic modeling unit of the system with transparency and voxel characteristics. The system describes various light spots, colors, and other light characteristics passing through the volume data field under certain lighting conditions and uses this to represent the internal information of the data field.

In traditional visualization models, we simply describe the data, but it is not possible to visualize the problem itself. Therefore, in order to better represent the mathematical ideas and methods applied to real-life problem solving, firstly, a solid diagram can be drawn to realize the function of software technology analysis; secondly, it can be converted into images, text, and other forms in computer language and expressed to the user through the combination of graphics and text; finally, the model can be processed on the computer using visualization tools to get the final result.

According to the different rendering order, the volume rendering algorithm can be divided into the volume rendering algorithm in the order of the image space and the volume rendering algorithm in the order of the object space. There are also certain differences between the two. The volume rendering algorithm occupies the image space from each pixel on the screen and emits light according to the direction of the view. These rays pass through a three-dimensional data field, and the same sampling is performed on the rays to find the sampling point. These rays pass through a three-dimensional data field, and the same sampling is performed on the rays to find the sampling point. The volume rendering algorithm occupies the object space according to a specific projection area and projection direction. Starting from a three-dimensional data field, project all the frames onto an image plane in a certain order, calculate the contribution of each data point, and obtain the final image on the screen. The typical representative of this type of algorithm is the method of throwing snowballs.

The commonly used method in the volume rendering algorithm described in this article is the ray casting algorithm. The principle of the algorithm is to decompose complex nonsimple objects into several small units; these units are called "blocks," that is, a basic unit. In the actual calculation, there are often some seemingly contradictory but interconnected units, and the relationship between these units is often very complex, so how to simplify the processing of these units has become one of the key issues in the visualization algorithm; the most common processing method is the local optimal solution method.

The main advantage of the ray casting method is that the use of gradients to change the boundary of the material can better reflect the normal direction of the surface of the object; using the phone model, specular reflection,

diffuse reflection, and environment reflection can achieve better lighting effects, and the sampling step can be adjusted to reflect local details.

*2.4. Showing.* The visualization display module is a function in the visualization-based student domain statistics software to provide data processing and graphical presentation. The system consists of three main parts: (1) raw data acquisition; (2) calculation and analysis of results output; (3) attribute assignment and weighting coefficients for input variables and other operations [14].

First, all information available within the scope of the defined original element is extracted, and then this information is converted into subblocks that represent the name, nature, and type of the data object in the visual display module, as needed. The subblocks representing the names of the data objects can be converted directly into the information contained in the visual display module, or the information can be encoded and then represented in numerical or pictorial form by a computer or software for the original elements' characteristic property variables and other related relationships.

*2.5. Application of Mathematical Models to Inputs and Outputs of Economic Variables.* The application of mathematical models in economic modeling is very broad and complex, and it is mainly used in the calculation and analysis of macroeconomics between factors of production, between capital and labor, and in the field of microeconomics. It is particularly important to use mathematical modeling to solve practical problems in some special cases.

*2.6. Mathematical Production Function Model.* A production function is a mathematical description of the quantitative relationship between factor inputs, product output, and technological progress in a production process. The production function can be used to describe a company or industrial production process. When studying macroeconomic problems, the production process can also be described by considering the whole economic system as a whole enterprise. It is widely used in economic theory research, production modeling, technical progress measurement, production capacity analysis, and economic forecasting. Since 1928, when the production function was proposed by researchers, economists have paid great attention to the production function. Production function models include Cobb-Douglas production function, constant elasticity of substitution production function, variable elasticity of substitution production function, and transcendental logarithmic production function [15].

The most widely used production function is the Cobb-Douglas production function whose general form is shown in the following formula:

$$Y = AK^\alpha L^\beta. \quad (4)$$

In the formula,  $Y$  represents output,  $K$  and  $L$  represent the input of capital and labor,  $A$ ,  $\alpha$ , and  $\beta$  are parameters and  $A > 0$ ,  $0 < \alpha < 1$ ,  $0 < \beta < 1$ . The Cobb-Douglas production function is generally called the C-D production function.

There are some improved forms of the Cobb-Douglas production function. One of them is the transcendental production function, whose form is shown in the following formula:

$$Y = AK^\alpha L^\beta e^{\alpha' L + \beta' L}, A > 0, \alpha', \beta' \leq 0. \quad (5)$$

If  $\alpha'$  and  $\beta'$  are equal to zero, the formula is simplified to the Cobb-Douglas production function form. Take the logarithm as shown in the following formula:

$$\ln Y = \ln A + \alpha \ln L + \beta \ln K + \alpha' L + \beta' K. \quad (6)$$

For this function, the marginal product may rise before it eventually falls. This function also allows output elasticity and substitution elasticity to vary with input changes.

The second improved form of the Cobb-Douglas production function is the Zellner-Revankar production function, whose form is shown in

$$Ye^{cY} = AL^\alpha K^\beta, c \geq 0. \quad (7)$$

Here, if  $c = 0$ , it is simplified to the form of Cobb-Douglas production function. Take the logarithm as shown in

$$\ln Y + cY = \ln A + \alpha \ln L + \beta \ln K. \quad (8)$$

The third improved form of the Cobb-Douglas production function is the Nerlove-Linstad production function, whose form is shown in

$$Y^{1+c \ln Y} = AL^\alpha K^\beta, c \geq 0. \quad (9)$$

Here, if  $c = 0$ , it is simplified to the form of Cobb-Douglas production function.

The fourth improved form of the Cobb-Douglas production function is the transcendental logarithmic production function, and its form is shown in

$$\ln Y = \ln A + \alpha \ln L + \beta \ln K + \gamma \ln L \ln K + \delta (\ln L)^2 + \varepsilon (\ln K)^2. \quad (10)$$

**2.7. MCD Mathematical Production Function Model for Analysis of Economic Production Function.** The production function is a mathematical description of the quantitative relationship between the input of production factors, product output, and technological progress in the production process. Production functions can be used to describe company or industrial production processes. When studying macroeconomic issues, the entire economic system can also be regarded as a whole enterprise to describe

the production process. It is widely used in the fields of economic theory research, production model establishment, technological progress measurement, production capacity analysis, and economic forecasting. Since the production function was proposed by researchers in 1928, economists have paid great attention to the production function. Production function models include Cobb-Douglas production function, fixed substitution elastic production function, variable substitution elastic production function, and transcendental logarithmic production function.

Since its inception, the Cobb-Douglas production function has been widely studied and applied by many economists due to its excellent economic properties and easy solution. The quantitative economic model derived from the Cobb-Douglas production function has been widely used in economic theory and production practice research. With the deepening of production and technology research and development, one-sided shortcomings have gradually appeared. The academic circles have never stopped arguing about this, and research on improvements and revisions will continue [16].

First of all, some basic properties of the Cobb-Douglas production function are very related to the actual production process. If the output elasticity does not change, the factor substitution elasticity is 1. For the same research object, if the sampling interval is different, the output elasticity and factor substitution elasticity should be different. Although the research objects are the same and the sampling interval is the same, for different sample data points, because the ratio of input factors to production factors is different, the elasticity of output and the elasticity of substitution must be different.

Second, this model is limited in structure. The relationship and structure of variables in the economic system are very complex, which leads to the fact that the Cobb-Douglas production function often cannot match historical data with satisfactory accuracy in practical applications. Reason analysis shows that, with the advancement of science and technology, the innovation of production equipment, the improvement of production department functions, the improvement of production efficiency, and the diversification of products, the production process of most enterprises or industries shows the characteristics of multiple input elements and diversified structures. The entire production process of most enterprises or industries may have subproduction processes, and these subproduction processes are mutually influencing and interrelated. When scholars need to use the production function to describe the total production output of a specific company or industry and the functional relationship between capital and labor input, they often use indicators such as average production elasticity and average technological progress. It is not difficult to see from the form of the Cobb-Douglas production function that it is suitable for describing the input-output relationship that represents a single production process. It is difficult to describe the complex production structure using Cobb-Douglas production function. This requires people to develop new production functions.

In actual economic production activities, the entire production process of a company or industry with a

diversified production structure consists of several sub-production units. Based on the previous discussion, coupled with the idea of additive nonparametric regression model, it is reasonable to assume that a production function showing this input-output relationship must be a combination of several basic production functions. That is, a model is a group of production subunits used to express each other's effects. Because the Cobb-Douglas production function has good economic properties, it is widely used by economists. This article assumes that a reasonable production function model should be a composite of several Cobb-Douglas production functions, expressed as

$$Y = F(K, L) = F(K^{\alpha_1} L^{\beta_1}, K^{\alpha_2} L^{\beta_2}, \dots, K^{\alpha_m} L^{\beta_m}). \quad (11)$$

Therefore, this paper constructs a production function with the characteristics of multiple production structure, which is used to describe the input-output function relationship of a certain industry and enterprise. The specific form is shown in

$$Y = \sum_{i=1}^m A_i K^{\alpha_i} L^{\beta_i}. \quad (12)$$

It is reversed from the digital form. This article calls this type of power exponential summation production function called MCD production function (Modified Cobb-Douglas Production Function). It can not only maintain some good properties of the Cobb-Douglas production function, but also improve the fitting accuracy of the production function to historical data. At the same time, it can also reflect the diversified structural characteristics of production to a certain extent [17].

The econometric form of the MCD production function is shown in

$$Y = \sum_{i=1}^m A_i K^{\alpha_i} L^{\beta_i} + \mu. \quad (13)$$

In order to make the expression concise, suppose the  $A_i$  in the MCD production function is expressed as

$$A = \sum_{i=1}^m A_i, \lambda_i = \frac{A_i}{A}. \quad (14)$$

The above hypothetical model (11) can be expressed as shown in the following formula:

$$Y = A \sum_{i=1}^m \lambda_i K^{\alpha_i} L^{\beta_i} + \mu. \quad (15)$$

### 3. Experimental Analysis and Results

**3.1. Input-Output Line Model.** Suppose that the economy of a country (or region) is divided into  $n$  sectors, and each sector produces one (or one type of) product, the total of intermediate use and final use in each department, minus transfers and imports. This result is equal to the total output of the department, as shown in

$$\sum_{j=1}^n x_{ij} + (r_i + p_i + q_i + c_i + s_i + d_i + e_i) - o_i - i_i = y_i, i = 1, 2, \dots, n. \quad (16)$$

Let  $z_i = r_i + p_i + q_i + c_i + s_i + d_i + e_i - o_i - i_i$ ; get the following formula:

$$\sum_{j=1}^n x_{ij} + z_i = y_i, \quad i = 1, 2, \dots, n. \quad (17)$$

The direct consumption coefficient is the core of the static input-output model. It can organically combine economic factors with technical factors, and it can combine technical factors for economic analysis. In a certain period of time, if the role of production technology and intermediate products does not change, it can be considered that  $a_{ij}$  is relatively stable. Substitute  $x_{ij} = a_{ij}x_j$  into formula (17) to obtain

$$\sum_{j=1}^n a_{ij}x_j + z_i = y_i, \quad i = 1, 2, \dots, n. \quad (18)$$

**3.2. Input-Output Column Model.** The production input of each department is composed of three parts, namely, intermediate input, fixed asset depreciation, and initial input. From this, a balance equation can be obtained as shown in

$$\sum_{i=1}^n x_{ij} + E_j + w_j + b_j + t_j + m_j = x_j, \quad j = 1, 2, \dots, n. \quad (19)$$

Marked as  $h_j = E_j + w_j + b_j + t_j + m_j$ , then formula (19) can be changed to

$$\sum_{i=1}^n x_{ij} + h_j = x_j, \quad j = 1, 2, \dots, n. \quad (20)$$

Write the above formula in matrix form; get the following formula:

$$A_c X + H = X, \quad (21)$$

where  $H = (h_1, h_2, \dots, h_m)^T$  is the value-added column vector, as shown in

$$A_c = \begin{pmatrix} \sum_{i=1}^n a_{i1} & & \\ & \ddots & \\ & & \sum_{i=1}^n a_{in} \end{pmatrix}. \quad (22)$$

**3.3. Revision Method of the Input-Output Direct Consumption Coefficient.** The direct consumption coefficient is the basis for calculating the absolute consumption coefficient and establishing the mathematical model. The accuracy of the consumption coefficient directly affects the overall balance and quality of the plan. In order to ensure the accuracy of the direct consumption coefficient, it is important to master the quality of basic statistical data. At the same time, it is necessary

to understand some factors that directly affect consumption changes before the coefficient can be corrected.

In the revision of the visual model, the variables defined in the input-output model can be directly referenced, or they can be transformed into a correction independent of the established functions and not dependent on other known information. However, this approach is flawed in practical applications. If the previous year's data is used as the reference object, the estimation results will be biased from the real situation; if the time series factors are taken into account in the next year and the next months' statistics, we get a constant that was not included in the previous calculation of the variables defined in the visualization model and therefore cannot be revised directly, but only as a variable [18].

The establishment of the input-output model should follow the following principles:

- (1) The economic variables and GDP per capita are linearly related over a certain period of time, and their trends are basically consistent
- (2) There is interdependence between economic phenomena and other factors for a specific unit in the field under study

The following issues should be noted when constructing a mathematical visualization and quantitative analysis framework. Firstly, the relationship between quantitative and qualitative analysis should be paid attention to when constructing the mathematical visualization model. Secondly, the mathematical model should reflect the scientific and artistic aspects as much as possible. Finally, the indicators should be selected, adjusted, and applied according to the actual situation to meet the requirements of the prudence principle.

Today the direct consumption factor method has undergone many changes and it is mainly divided into non-mathematical methods and mathematical methods. Nonmathematical methods are such as the actual survey method and the expert assessment method. There is a mathematical method: the most basic one is the RAS method.

The RAS correction method is a direct consumption coefficient correction method, which is essentially an iterative correction method. The basic assumption of the RAS correction method is that the direct consumption coefficient changes come from two aspects.

The RAS method was proposed by the researcher Richard Stone in the 1960s. This method is also known as the bilateral proportional method. The idea of this method is to use the total average annual usage as the control amount for rows, and the average total input in the target year as the control amount for columns. The intermediate input matrix lists the conditions for double reduction, full utilization of the target year, and full input in the target year. Estimates are made using conventional statistical data, which greatly reduces the cost of the input-output table.

The most basic mathematical method in the direct consumption factor method is the RAS method, which has many advantages. (1) Using RAS directly as the basic idea of the correction model can make the calculation results more intuitive and clear, and it does not require any complicated and tedious mathematical steps in practical operation. (2) The visualization method enables us to observe, analyze, and understand the data from more perspectives.

#### 4. Results

- (1) The input-output model can be used to study many important proportional relationships in the national economy, which can be used to measure issues such as production or technological progress. For example, it can study the rate of intermediate products and the rate of final products, the value composition of products in a certain department, and so on.
- (2) Study the impact of changes in each final demand on the output value of each department, labor remuneration, and social net income, for example, the increase and decrease of the military budget, the expansion and compression of the scale of infrastructure construction, and the impact of the increase and decrease of imports and exports on the national economy.
- (3) Analyze the impact of changes in wages and taxes on the prices of products in each department.  $\Delta P$ ,  $\Delta V$ , and  $\Delta M$  are used to represent the column vectors composed of changes in prices, wages, and profits and taxes, respectively, as shown in

$$\Delta P = (I - A^T)^{-1} \Delta V, \Delta P = (I - A^T)^{-1} \Delta M. \quad (23)$$

- (4) Analyze the impact of changes in the price of a product on the prices of other products.
- (5) Conduct energy consumption analysis. Use the input-output table to calculate A and B, and further calculate the energy consumption indicators of the unit output value of the main sectors of the national economy. This is an important technical indicator, from which we can see the technical content and energy waste of products in the leading sectors of our national economy. It is helpful to find the gap and reduce the comprehensive energy consumption per unit of product.

#### 5. Conclusions

In short, the application of scientific computing visualization technology can greatly accelerate data processing, and it can realize image communication between people and computers. It also enables scientists to guide and control the calculation process and consciously control scientific calculations so that they can run. The development of scientific



## Research Article

# Multiobjective Fusion Decision Model for Key Technologies of Aircraft Intelligent Lightweight

Anliang Zhou <sup>1</sup>, Lei Dong <sup>2</sup>, Yuedong Lang <sup>3</sup>, Jiangping Cui <sup>4</sup>, Wei Ren <sup>5</sup>,  
Jiankang Liu <sup>2</sup> and Zhiqiang Wang <sup>2</sup>

<sup>1</sup>School of Mechanical and Electrical Engineering, Beijing Institute of Graphic Communication, Beijing 102600, China

<sup>2</sup>School of Mechanical Engineering, Tianjin University of Technology and Education, Tianjin 300222, China

<sup>3</sup>Beijing Aerospace Control Instrument Research Institute, Beijing 100854, China

<sup>4</sup>China Mobile Procurement Shared Service Center, Beijing 102200, China

<sup>5</sup>Beijing Readir Technology Co., Ltd, Beijing 100031, China

Correspondence should be addressed to Lei Dong; [leid@tute.edu.cn](mailto:leid@tute.edu.cn)

Received 30 December 2021; Accepted 21 January 2022; Published 15 February 2022

Academic Editor: Daqing Gong

Copyright © 2022 Anliang Zhou et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

As a tool for aviation exploration, aircraft plays a pivotal role in the aerospace field. The aircraft's lightweight can ensure that the aircraft can carry more loads, so it is of great significance to the research on the key technologies of aircraft lightweight. This paper organizes the intelligent technology system from the multiobjective fusion technology relationships and considers the complex changes between technology units. A technology group model for the intelligent technology system of spacecraft welding is formed to provide a decision-making basis for the realization of batch, stable, and efficient intelligent and lightweight manufacturing of spacecraft. We expand the technology development research from single technology research to technology system research. We research and build a three-dimensional perspective intelligent technology system based on technology unit, technology system, and technology group. It reveals the internal relationship description of technological development and expands and perfects the current research theory of technological development.

## 1. Introduction

Aircraft has been widely used and plays a pivotal role in the aerospace field [1–3]. As one of the key technologies in aircraft research and development, lightweight design has received extensive attention [4, 5]. Currently, intelligent technologies related to lightweight are being deeply integrated with traditional manufacturing processes and accelerate the conventional manufacturing process and development mode [6, 7]. It is triggering a new round in lightweight technology and industrial development models [8]. This change is reflected in increasingly lightweight designs that entirely apply intelligent technology to achieve an all-round digital and intelligent transformation of the manufacturing process, research and development, production, management, and service [9]. The quality of aircraft products and the production efficiency have been

continuously improved. It has driven the rapid development of many related intelligent technologies, intelligent equipment, and intelligent products, providing necessary manufacturing process support for the manufacturing industry [10, 11]. Therefore, the integration of lightweight and intelligent technology is of great significance for the development of aircraft intelligent manufacturing.

In recent years, many scholars have done a lot of research on lightweight technology. To sum up, it mainly includes the following three aspects: lightweight design technology, lightweight material technology, and lightweight manufacturing technology. Wu et al. [12] and Fiedler et al. [6] achieved the effect of lightweight design through topology optimization design under the premise of satisfying safety. Zuo [7] carried out structural optimization under the constraints of static stiffness and dynamic frequency stiffness and achieved good lightweight effect. Cho et al. [13] used the size optimization method to

optimize the design of the rail transit car body and achieved a lightweight effect by reducing the body weight by 29%. Hou et al. [14] optimized the multiobjective design of structural performance, cost, and mass and reduced the mass of the components while increasing the structural stiffness. Emirci and Yildiz [15] compared and analyzed the anticollision and energy-absorbing characteristics of the body energy-absorbing box using aluminium alloy and high-strength steel. Through lightweight material technology, the aluminium-made energy-absorbing box is better than the energy-absorbing effect per unit mass and the target of the steel energy-absorbing pack. Parsa and Darbandi [16] analyzed the application of hydroforming in car body forming and achieved lightening from the manufacturing technology.

Although the above lightweight methods have achieved good results, the applied methods are relatively single, and the influence of the mutual coupling of multiple factors has not been considered. For aircraft, various factors such as product quality, performance, and manufacturing cycle require higher lightweight technology [17]. Especially in the welding process of lightweight materials, it is urgent to change the current situation of unstable quality of welding products and dependence on operator skills [18–20]. The application of intelligent digital technology and lightweight knowledge base can realize the optimization of welding process parameters and the identification of welding defects [21, 22]. At the same time, uncontrollable factors such as welding cracks and deformation are eliminated, welding accuracy is improved, the welding production mode is changed, and high-quality and high-efficiency manufacturing of lightweight material welded products is supported [23]. To better combine intelligent technology with lightweight goal, it is necessary to clarify the correlation factors between lightweight technology and system analysis method. At the same time, a set of technical systems is formed according to its technical units, and a multiobjective fusion decision-making model is established. It will be of great significance to describe the development of technology units in the way of technology group [24–26].

This paper presents a multiobjective fusion decision model applied to the key technology of spacecraft. The model fully considers the complex changes between technical units, organizes the intelligent technology system, and forms a technology group for spacecraft intelligent lightweight technology system. The research on the development of lightweight technology is expanded from single technology research to the technical system of the connection between technical units. The three-dimensional perspective intelligent technology system of technology unit, technology system, and technology group is constructed. It provides a decision-making basis for spacecraft batch, stable, efficient, and lightweight manufacturing.

## 2. Concept and Method

*2.1. Mathematical Representation of the Technical System.* The technology system construction method aims to reveal the specific relationship research object of the intelligent technology unit set to the technology system with specific

relationships (defined as). The technology group (defined as) can be described as a set of specific relationships. The corresponding representations of the three are as follows:

Technical unit set:  $Y = \{1, 2, 3, \dots, y\}$

Technical system set:  $P = \{Y_1, Y_m, Y_n, \dots\}$

Technology cluster:  $T = \{T_1, T_m, T_n, \dots\}$ ,  $T$  for a specific connection

The system construction of technology development can be described as defining a specific relationship  $T$  on the set  $Y$ , which describes the specific relationship from the technology unit set  $Y$  to the case set  $S$ , which is described as  $T \subseteq Y \times S$ . That is, the relationship  $T = \{Y, T_C\}$  is defined on the technical unit set and the relationship  $T = \{S, T_Q\}$  is defined on the case set, so  $T$  describes  $T$ , respectively,  $T_C$  and  $T_Q$  define two specific relation operations on sets  $Y$  and  $S$ , so that a group of things with regular interaction or interdependence form a whole, that is, the technical system  $P$ ,  $P$  is described as the abstract system subset  $P: T \subseteq Y \times S \rightarrow Y$  formed by the relation operation from the technical unit set  $Y$  to the specific relation  $t$  of the case set  $S \times S \rightarrow Y$ . The specific relationship of  $S$  is the mapping from  $T$  to  $Y$ . Furthermore, the technical system  $P$  is described as  $Y, \times, S$ . That is, the technical system  $P$  is a subset of the specific relationship  $T$  from the technical unit set  $Y$  to the case set  $S$ .

*2.2. Unit Analysis of the Technical System.* Make a good selection of technical units with specific internal relations to form a technical system around internal ties. Several technical systems form a technical group with specific internal ties. This method includes a dynamically related technical unit technical system-technical group to describe the relationship between the development of intelligent technology. Based on technical units, the technology group and technology system for decision-makers can be optimized, and the technology system can be adjusted and optimized through the technology group, or the new technical units can be expanded and supplemented to form a set of dynamic evolution correspondence. The technical system composes technical units, and the technical units are continuously developed and evolved. For a specific technical unit of a system, it will not only be affected by other technical units in the system but also by changes in the external environment and be reorganized with changes in the objectives of the technical system, as shown in Figure 1.

Y1 The technical complexity of Y2 and Y3 is different. For example, the technical system composed of special alloy precision casting technology, comprehensive design optimization technology, and virtual manufacturing technology based on product life cycle management describes the technical system of special alloy precision casting facing the full life cycle of products with different dimensions of process, process design and process production management, and the comprehensive design optimization technology refers to special alloy. In the design stage of gold casting products, the technology or technical complexity required is different from the whole life cycle management of precision casting equipment

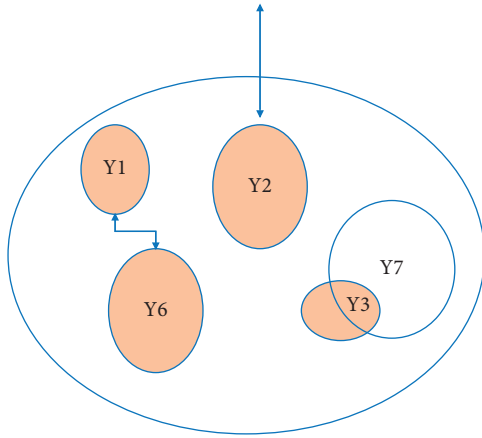


FIGURE 1: Evolution of technology unit relationship.

and the manufacturing process. Similarly, the coverage and complexity of special alloy precision manufacturing equipment technology, comprehensive design optimization technology, and life cycle management technology of manufacturing process are also different. At the same time, the comprehensive design optimization technology has many common applications, which can be used in other technical systems outside the technical system and can also improve its design optimization ability and methods due to the changes of the external development environment, the comprehensive optimization design technology itself also needs to consider the ability of special alloy precision casting technology. The comprehensive optimization design beyond the capability is not supported by special process equipment for equal material manufacturing. The two are combined, as shown in Figure units Y3 and Y7. Product life cycle management technology can be used both inside and outside the technical system. As shown in Figure 1, Y2 technical units will also be greatly affected by changes in technical means or environment outside the technical system.

The specific relationship between Y1 and Y6 and Y3 and Y7 described in the technical unit also has the difference of sequential relationship and cross-relationship. For example, the technical system is composed of welding process real-time monitoring and detection technology, welding defect automatic identification technology, high-precision welding robot, and welding power supply. The welding process real-time monitoring and detection technology are the premise of welding defect automatic identification technology. With the welding process real-time monitoring and detection technology, the welding defect automatic identification technology can be carried out. The two restrict each other, appear in a sequence, and work together, as shown in Y1 and Y6 in Figure 2. The high-precision welding robot and welding power supply are closely connected, overlapped, and crossed to form an association relationship between connecting electric energy and forming a stable execution action, as shown in Y3 and Y7 in Figure 1. The diagram only shows the order and cross between technical units, and there may be more connections.

Y2 described in the technical unit interacts with the outside of the technical system, which may be affected by the external technical system or environment or may affect the external technical system or environment. For example, the welding expert system Y2 will not only affect the welding production system but also affect the welding production process layout. In addition, the technical system is also affected by process objectives or value objectives, and even the changes of objectives and systems caused by changes in the external environment should be considered. For example, there are differences in the correlation model of technical units of a welding system aiming at high precision or high efficiency; for another example, for a certain product, whether the market is easy to accept personalized customization or more inclined to mass manufacturing also has a great impact on the technical system.

**2.3. Correlation Analysis of the Technical System.** The technology system is composed of technology units with specific association relations. Different association relations have a great impact on the technology system. There are many kinds of association relationships. For example, there is a relationship between process and function between product process and intelligent function. There is a relationship between product quality improvement and production line production efficiency, stability, and consistency, as well as process optimization and automation improvement. As shown in Table 1, List T1 (production line production efficiency), T2 (intelligent smelting system), and T3 related to its technical unit (equipment fault prediction and diagnosis) is a specific correlation. Based on the listed correlations, the technical unit forms a specific target-oriented technical system or technology group. It describes the technical systems of the four types of correlation on the right side of the table. Due to the correlation, the orderly combination of multiple technical units is realized to form the technical system described in the table. The correlation of the technical system is the decisive factor of organizing the technology units and is also the goal of the formation of the technology system.

**2.4. Technical System Construction of Multiobjective Fusion.** In the construction of the technical systems, we often consider the fusion scheme aiming at a variety of correlation relations and comprehensively consider the complex changes between technical units to form a multiobjective fusion technical system [27–29]. It is precise because of the complexity of the integration of multiple relationships and the complexity of the relationship between technical units that constitute the complexity and uncertainty of the technical system. The technical systems constructed are different due to different technical associations T1 and T2; due to other fusion targets TM and TN, other technical systems are built, as shown in Figure 2.

Figure 2 reveals the complexity and diversity of intelligent technology systems. It can be seen from the diagram that the cluster  $T = \{T_1, T_m, T_n, \dots\}$ . The research of technology connection based on the technology unit library has

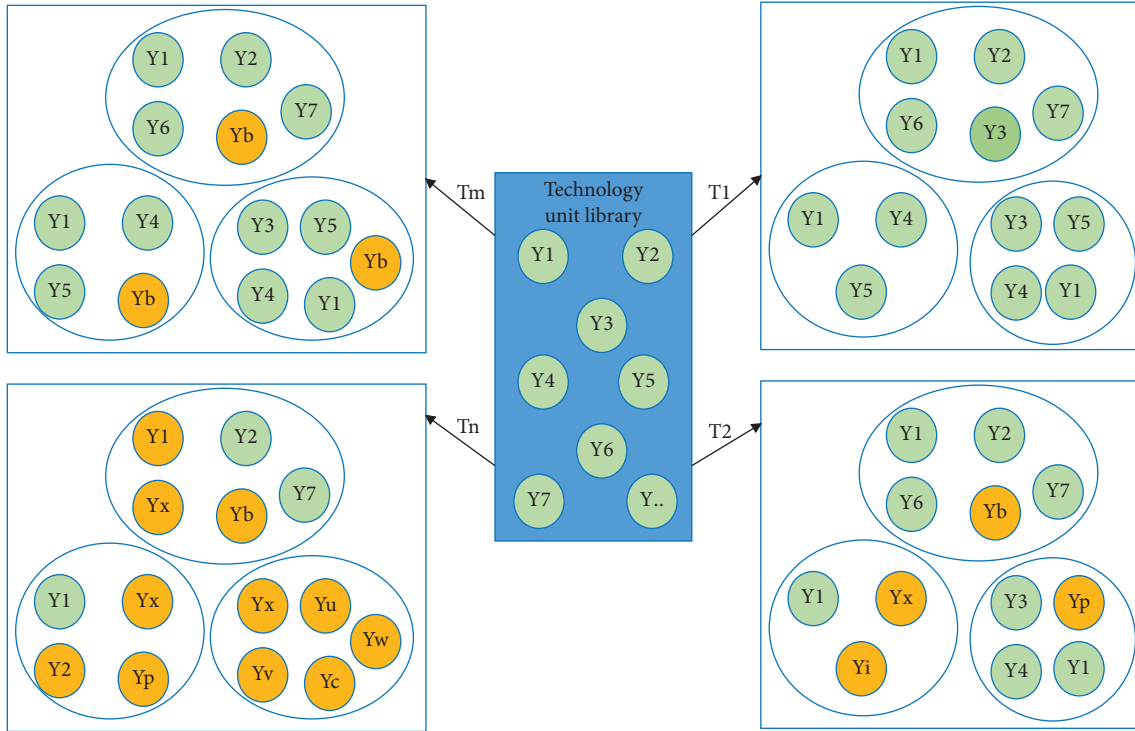


FIGURE 2: Technical system construction diagram of multitarget fusion.

TABLE 1: Example of system association relationship construction.

Number	Association relationship	Technical system example
T1	Production efficiency of the welding production line	Intelligent welding process control: welding seam forming control technology based on visual sensing, constant pressure adaptive control technology, welding path adaptive control technology, and arc length adaptive control technology based on laser sensor tracking. These four central technical units constitute a technical system for realizing the automatic control of welding path and welding process parameters and improving the automatic control level of the welding process.
T2	Casting melting intelligent system	Casting intelligent melting system: it is composed of intelligent feeding and unloading technology, online detection technology of molten iron parameters, LIMS (Laboratory) information management technology, and intelligent 3D printing technology of casting mould. These four main technical units constitute the technical system to cast the intelligent melting.
T3	Fault prediction and diagnosis of forging equipment	Equipment fault prediction and diagnosis: it comprises equipment fault diagnosis system technology, product data acquisition technology using RFID, QR code/bar code, intelligent control, intelligent decision support, and visualization tool technology. These three main technical units aim at equipment fault prediction and diagnosis.

become an eternal theme. It can also be considered that the research on the development of intelligent technology focuses on the change of technology units and technology connections. The change of technology units, technology connection, and the evolution of external environment of technology system constitute a complete system of technology development.

### 3. Multiobjective Fusion Decision-Making Model

Because of the above specific problems of the aircraft welding process, the key to the development of intelligent

technology lies in the intelligent improvement of the quality, performance, and manufacturing cycle of relevant products, that is, the target relationship of aircraft lightweight material welding technology system. The traditional lightweight material welding workshop has some prominent problems. Product quality control does not form a closed loop, process design depends on human experience, low degree of equipment automation, backward detection means, and so on. There is an urgent need to transform and upgrade to the intelligent direction to achieve the synchronous improvement of product consistency, manufacturing flexibility, and efficiency. A multiobjective fusion relationship has been formed for such problems, then selecting relevant



technology units from the Technology Library and external technology environment, as shown in Figure 3.

Based on the final goal and demand of lightweight material welding, combined with the technical unit library, seven intelligent technology systems  $P = \{Y_1, Y_2, \dots, Y_7\}$  for the intelligentization of lightweight material welding workshops are analyzed and the following seven major process problems are solved, respectively:

$P = \{$

Question 1: describe the welding method, process parameters, and scope of the configuration of the welding equipment, drive, and control system.

Question 2: establish a database to store and compile the data of various existing welding production processes for sharing by all welding workstations.

Question 3: according to the welding method, welding material types, and specifications and other data, formulate a program for optimizing welding process parameters.

Question 4: according to the workpiece shape, size, and the preset deviation limit of the seam, the automatic correction and compensation program are compiled.

Question 5: according to the image of the welding area, according to the shape of the welding arc and the welding bead, the image recognition, and remote monitoring program are compiled.

Question 6: prepare the automatic fault diagnosis, alarm, and repair program for the drive system and control system of welding equipment, wire feeder, and welding power source.

Question 7: manage, distribute, and collect the data between the welding equipment control system and the DCS control system operating station and realize the monitoring and remote diagnosis functions of the processing process.

$\}$

The corresponding eight technical systems are as follows:

Y1 refers to rapid interchange clamping (flexibility). Y2 refers to the instrumentalization of weld quality inspection. Y3 refers to the digitization of welding production management. Y4 refers to the real-time management of welding control. Y5 refers to the paperless basic welding database. Y6 refers to the computerization of weldability analysis. Y7 refers to the intellectualization of welding process analysis.

From the dimensions of different fusion associations, the decision model for the association between technical units-technical systems-technical groups is described as follows.

From the shop floor construction and target demand dimension. DCS unified the control and deployment of the intelligent welding workshop, the production tasks issued by the management, the manufacturing unit prefabricated action program, and test data according to the instructions to determine the entire manufacturing process. Everything is ready, the work program starts the welding command, and the main control unit collects the processing parameters in

real time. At the same time, online inspection, quality control, welding process monitoring, analysis, and process optimization and adjustment are carried out from the dimension of welding technology development. Intelligent welding technology effectively improves the accuracy and reliability of product quality. Focusing on the continuous improvement of welding control technology such as welding quality, welding efficiency, and welding spatter, foreign manufacturers have taken the lead in realizing the application of full digitalization in welding equipment, including the precise operation control and working process control of the microprocessor to achieve various welding performances. These digital and intelligent applications make welding integrated and simplified, no more reliance on welders and traditional welding processes. At the same time, in a concise time, the change of arc length can be controlled accurately and reliably. The software upgrade of the welding equipment can also be applied to the needs of different occasions to ensure the beauty of the welding seam. Due to the precise computing capability of the microprocessor, various welding specifications and precise arc starting and ending can be achieved.

Therefore, the aircraft welding workshop is applied to seven technical systems through the decision model with the dimensions of different correlation and technology development, integration, quick interchange clamping (flexibility), totalization of weld quality inspection, vitalization of welding production management, real-time welding control management, closed-loop welding process quality control, paperless welding basic database, datalization of weldability analysis, and intellectualization of welding process analysis. It is precisely because of the application of the technology group formed by the construction of these eight technical systems that the batch, stable, and efficient manufacturing of aircraft is realized manufacturing capacity.

#### 4. Results and Discussion

To further improve the production capacity and technical level of the welding process of aircraft lightweight materials, China has carried out a series of intelligent technology innovations for domestic aircraft manufacturing enterprises with the support of 04 unique project. The technical means adopted for the welding process of lightweight aircraft materials are shown in Figure 4.

The main problems existing in the welding of lightweight aircraft materials can be summarized into the trial set to be solved by the technology group, that is, the correlation set:

$T = \{$

T1: traditional methods based on manual and experience are primarily used in auxiliary links such as clamping and measurement, automatic welding, parameter selection, and process control. There are still difficulties of insufficient product quality stability and low efficiency.

T2: product quality depends on the skills of skilled workers, and there is no closed loop of product quality

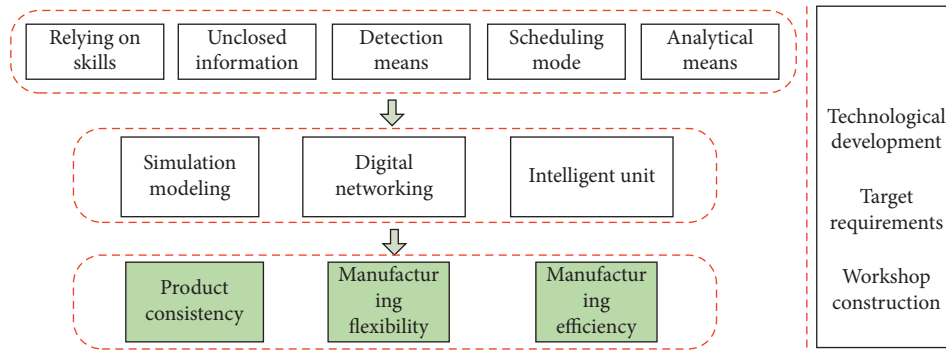


FIGURE 3: Element Atlas of lightweight material welding workshop.

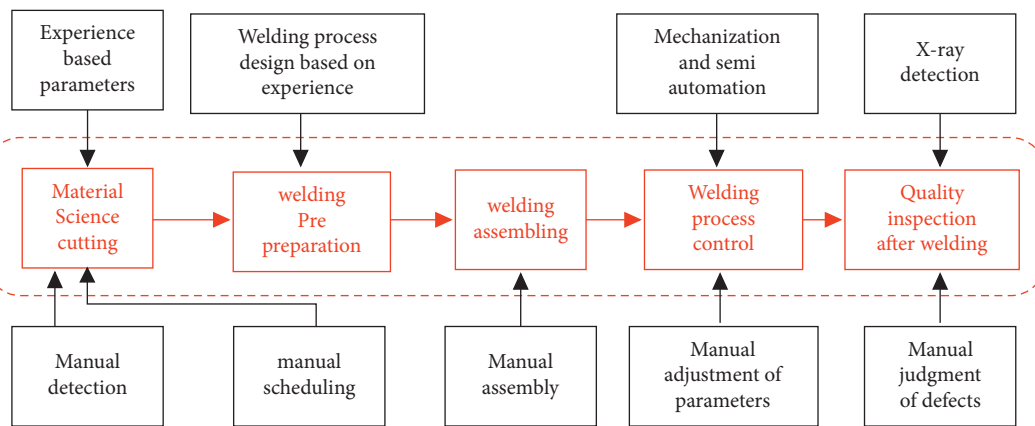


FIGURE 4: Manufacturing technology means of the current welding workshop.

control, and the stability of product quality is difficult to control.

T3: restricted by the relatively backward parameter, optimization level and detection methods, defect determination, scheduling method, clamping methods, and other factors, the demand for fast and efficient manufacturing is difficult to achieve.

T4: product manufacturability analysis still relies on personal knowledge and experience. The flexibility of product welding and manufacturing and the ability to quickly replicate the production line cannot meet the actual needs.

}

First, the leading technical units in lightweight material welding are analyzed. Manufacturing enterprises in automobiles and airplanes in industrialized countries have widely applied automation technology, sensor technology, intelligent technology, information technology, robot technology, etc., and gradually realized the digitization and intelligence of manufacturing workshops, representing a higher level of development of manufacturing workshops. Among them, the more typical technical groups influencing factors for the development of technical units are the following: (1) The rapid development of products or equipment is realized by digital modelling and simulation of products

and workshops, and some products or components have been successfully developed, shortening the product development cycle to achieve rapid response manufacturing. (2) Through the upgrade of the digital network of the manufacturing site, to realize the monitoring of equipment operating parameters in the production process of the enterprise and the process monitoring of production system failures. (3) Realize workshop intelligence with the precision and intelligence of manufacturing equipment units.

Second, combined with the relationship of multi-objective fusion of aircraft, the improvement of welding manufacturing efficiency, manufacturing flexibility, and product consistency of lightweight materials can be achieved. The intelligent welding workshop adopts advanced technical means such as intelligent manufacturing unit, simulation modelling, and digital networking. To solve various existing problems. Through the intelligent upgrade, the elements of technology development dimension, target demand dimension, and workshop construction dimension are satisfied. It is also the influencing factor of the technology group in the development of the technology unit.

Finally, the influence of technology group development is analyzed from the perspective of welding technology and process function. The technical units of the unit library of multiple welding workshops are selected for analysis. Taking a typical welding workshop as an example, its process and



corresponding technical units are described as follows: (1) Define the welding method, process parameters, and scope. (2) A database is established to store the data of various welding production processes and shared by all welding workstations. (3) According to the data of welding method, welding material type and specifications, etc., formulate a program for optimizing welding process parameters. (4) According to the workpiece shape, size, and the preset deviation limit of the seam, the automatic correction and compensation program are compiled. (5) According to the real-time recording of the welding area image during the welding process, the remote monitoring program is compiled according to the welding arc and welding beam shape parameters. (6) Set up automatic fault diagnosis alarm and repair procedures, which are applied to the driving system and control system of welding equipment, as well as wire feeder and welding power supply. (7) Manage, distribute, and collect the data between the welding equipment control system and the DCS control system operating station and realize the monitoring and remote diagnosis functions of the processing process.

## 5. Conclusions

This article proposes a multiobjective fusion decision-making model for aircraft intelligent lightweight key technologies. The model fully considers the complex changes between technical units and describes the development of technical units in a technical system in the form of technical groups. Extend the lightweight technology from single technology research to a technology system that studies the connection between technology units. In this paper, a technical unit technical system-technical group three-dimensional perspective intelligent technology system is constructed. It reveals the description of the internal connection of technology development and expands and perfects the current research theory of technology development. And it provides a decision-making basis for the batch, stable, efficient, and lightweight manufacturing of spacecraft. At the same time, the model can further expand the scope of industrial applications, which is very beneficial for improving the decision-making method of various equipment and systems' lightweight and provides a powerful tool for the lightweight design of various equipment.

To sum up, this paper studied the multiobjective of the key technology of aircraft intelligent lightweight and proposed a multiobjective fusion decision-making model. Due to the limitations of experimental conditions, time, and energy, the model needs to be further improved: first, the key technologies of intelligent lightweight cover a wide range, and the factors are influenced and coupled with each other, and it is difficult to cover the expert knowledge fully. The selected experimental cases and technical analysis are inevitably one-sided and limited, so it is necessary to improve the model further. Second, due to the small number of aircraft samples, the selected technical units, technical systems, and technical groups have limitations in the stereoscopic perspective. Further analysis, mining, and improvement are needed to increase the scope of application.

## Data Availability

The data used to support the findings of this study are included within the article.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References

- [1] M. Abi and E. Simic, "Airport categorization and possible improvements in accordance with icao global air navigation plan 2016-2030," *Traffic Engineering & Communications*, vol. 1, 2019.
- [2] Eurocontrol, *Air Traffic Flow and Capacity Management Operations Manual-Network Manager*, European Organisation for the Safety of Air Navigation, Brussels, Belgium, 2020.
- [3] Y. Xu, X. Prats, and D. Delahaye, "Synchronised demand-capacity balancing in collaborative air traffic flow management," *Transportation Research Part C: Emerging Technologies*, vol. 114, pp. 359–376, 2020.
- [4] G. W. Jang, M. S. Yoon, and J. H. Park, "Lightweight flatbed trailer design by using topology and thickness optimization," *Structural and Multidisciplinary Optimization*, vol. 41, no. 2, pp. 295–307, 2010.
- [5] N. Tanlak, F. O. Sonmez, and M. Senaltun, "Shape optimization of bumper beams under high-velocity impact loads," *Engineering Structures*, vol. 95, pp. 49–60, 2015.
- [6] K. Fiedler, B. F. Rolfe, and T. De Souza, "Integrated shape and topology optimization-applications in automotive design and manufacturing," *SAE International Journal of Materials and Manufacturing*, vol. 10, no. 3, pp. 385–394, 2017.
- [7] W. Zuo, "Bi-level optimization for the cross-sectional shape of a thin-walled car body frame with static stiffness and dynamic frequency stiffness constraints," *Proceedings of the Institution of Mechanical Engineers - Part D: Journal of Automobile Engineering*, vol. 229, no. 8, pp. 1046–1059, 2015.
- [8] S. Hunkeler, F. Duddeck, M. Rayamajhi, and H. Zimmer, "Shape optimisation for crashworthiness followed by a robustness analysis with respect to shape variables," *Structural and Multidisciplinary Optimization*, vol. 48, no. 2, pp. 367–378, 2013.
- [9] L. Duan, N. Xiao, Z. Hu, G. Li, and A. Cheng, "An efficient lightweight design strategy for body-in-white based on implicit parameterization technique," *Structural and Multidisciplinary Optimization*, vol. 55, no. 5, pp. 1927–1943, 2017.
- [10] X. Li, "Study on the progress of welding science and technology in China," *Journal of Mechanical Engineering*, vol. 48, no. 6, pp. 19–31, 2012.
- [11] D. Lee, N. Ku, T. W. Kim, J. Kim, K.-Y. Lee, and Y.-S. Son, "Development and application of an intelligent welding robot system for shipbuilding," *Robotics and Computer-Integrated Manufacturing*, vol. 27, no. 2, pp. 377–388, 2011.
- [12] J. Wu, S. Zhang, Z. Xie et al., "Topology optimization and analysis for frame structure of FSAE racing car," *Automobile Applied Technology*, vol. 6, pp. 115–119, 2018.
- [13] J. G. Cho, J. S. Koo, and H. S. Jung, "A lightweight design approach for an EMU carbody using a material selection method and size optimization," *Journal of Mechanical Science and Technology*, vol. 30, no. 2, pp. 673–681, 2016.
- [14] W. Hou, Z. Wang, W. Zhang et al., "Optimization design for auto-body beam section based on complex engineering

- constraints,” *Journal of Mechanical Engineering*, vol. 50, no. 18, pp. 127–133, 2014.
- [15] E. D. emirci and A. R. Yildiz, “An investigation of the crash performance of magnesium, aluminum and advanced high strength steels and different cross-sections for vehicle thin-walled energy absorbers[J]. Materialpruefung: werkstoffe und Bauteile,” *Forschung Prufung Anwendung*, vol. 60, no. 7-8, pp. 661–668, 2018.
- [16] M. H. Parsa and P. Darbandi, “Experimental and numerical analyses of sheet hydroforming process for production of an automobile body part,” *Journal of Materials Processing Technology*, vol. 198, no. 1-3, pp. 381–390, 2008.
- [17] A. K. Singh, V. Dey, and R. N. Rai, “Techniques to improve weld penetration in TIG welding (A review),” *Materials Today Proceedings*, vol. 4, no. 2, pp. 1252–1259, 2017.
- [18] G. C. M. Patel, P. Krishna, and M. B. Parappagoudar, “An intelligent system for squeeze casting process-soft computing based approach,” *International Journal of Advanced Manufacturing Technology*, vol. 86, no. 9-12, pp. 3051–3065, 2016.
- [19] W. Tu, X. Zhang, H. Shen, and B. Liu, “Numerical simulation on multiple pouring process for a 292 t steel ingot,” *China Foundry*, vol. 11, no. 1, pp. 52–58, 2014.
- [20] Y. Tanaka and I. Sato, “Development of high purity large forgings for nuclear power plants,” *Journal of Nuclear Materials*, vol. 417, no. 1-3, 2011.
- [21] G. C. M. Patel, P. Krishna, and M. B. Parappagoudar, “An intelligent system for squeeze casting process-soft computing based approach,” *International Journal of Advanced Manufacturing Technology*, vol. 86, no. 9-12, pp. 3051–3065, 2016.
- [22] O. S. Logunova, I. I. Matsko, I. A. Posohov, and S. I. Luk’ynov, “Automatic system for intelligent support of continuous cast billet production control processes[J],” *International Journal of Advanced Manufacturing Technology*, vol. 74, no. 9-12, pp. 1407–1418, 2014.
- [23] P. Freytag, “The renaissance of hydroforming-European market development, challenges and technical responses from the perspective of a serial manufacturer,” in *Proceedings of the 8th Int. Conf. on Hydroforming*, pp. 20–28, Stuttgart, Germany, 2014.
- [24] B. Yoon and B. Song, “A systematic approach of partner selection for open innovation,” *Industrial Management & Data Systems*, vol. 114, no. 7, pp. 1068–1093, 2014.
- [25] Y. Jeong, I. Park, and B. Yoon, “Forecasting technology substitution based on hazard function,” *Technological Forecasting and Social Change*, vol. 10, no. 1, pp. 259–272, 2016.
- [26] R. Kapoor and J. M. Lee, “Coordinating and competing in ecosystems: how organizational forms shape new technology investments,” *Strategic Management Journal*, vol. 34, no. 3, pp. 274–296, 2013.
- [27] K. Adamus, Z. Kucharczyk, K. Wojsyk, and K. Kudla, “Numerical analysis of electron beam welding of different grade titanium sheets,” *Computational Materials Science*, vol. 77, no. 5, pp. 286–294, 2013.
- [28] T. Kovacs, “Laser welding process specification base on welding theories,” *Procedia Manufacturing*, vol. 22, no. 3, pp. 147–153, 2018.
- [29] S. Pattnaik, D. B. Karunakar, and P. K. Jha, “Developments in investment casting process—a review,” *Journal of Materials Processing Technology*, vol. 212, no. 11, pp. 2332–2348, 2012.

## *Retraction*

# **Retracted: A Decision Method for Benefit Distribution Mechanism of Shared End Distribution on Shapley Value**

### **Computational Intelligence and Neuroscience**

Received 17 October 2023; Accepted 17 October 2023; Published 18 October 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### **References**

- [1] H. Zhang, Y. Hou, and F. He, "A Decision Method for Benefit Distribution Mechanism of Shared End Distribution on Shapley Value," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3816440, 13 pages, 2022.

## Research Article

# A Decision Method for Benefit Distribution Mechanism of Shared End Distribution on Shapley Value

Hao Zhang <sup>1</sup>, Yuan Hou <sup>2</sup>, and Feng-Feng He <sup>3</sup>

<sup>1</sup>Beijing Technology and Business University, Beijing Food Safety Research Base, Beijing, China

<sup>2</sup>Beijing Technology and Business University, Beijing, China

<sup>3</sup>Southwesten University of Finance and Economics, Chengdu, China

Correspondence should be addressed to Hao Zhang; zhaozhao@126.com

Received 5 January 2022; Accepted 28 January 2022; Published 14 February 2022

Academic Editor: Daqing Gong

Copyright © 2022 Hao Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Online shopping has led to the rapid development of e-commerce, and at the same time, the pressure of offline distribution has increased abruptly. Therefore, a current development trend is to share end-to-end distribution against the background of the Internet. The main research content of this paper is the benefit distribution mechanism of shared end distribution. Based on an analysis of the current situation of interest distribution, this paper proposes factors that affect interest distribution from the perspectives of individuals and groups. The suitable income distribution mode of enterprise alliances is chosen from two dimensions—cooperation mode and coordination mechanism. Based on extant theory, this paper proposes a benefit distribution scheme-selection mechanism based on the modified Shapley value method and takes the terminal distribution in the Haidian District of Beijing as an example. The revised income distribution results better reflect the income-generating abilities of different enterprises within a cooperative organization and assign different benefit proportions to this cooperative organization based on different income-generating capacities to provide development incentives and, at the same time, better achieve income distribution.

## 1. Introduction

The terminal distribution cost of logistics enterprises remains constantly high; the key problem is that the dependence on human resources limits the ability of these enterprises to reduce human costs, and the allocation of logistics facilities resources is unreasonable [1–3]. The logic behind “Internet+logistics” is not applied to solve the problem of low-efficiency logistics operations. For example, cost waste is caused by an invalid operation caused by circuitous transportation, no-load transportation, and secondary distribution. To eliminate the dependence on human resources, improve the allocation efficiency of logistics facilities resources, and reduce the chances of invalid operations, it is necessary to make the end distribution operations automated, large-scale, systematic, and intelligent. Shared end distribution means that a number of logistics enterprises participate in the end distribution, that is, through

connecting transfer points, order consolidation, and other ways to integrate end distribution infrastructure resources to provide low-cost and efficient end distribution services [4, 5]. Shared end distribution through horizontal logistics integration, that is, the unified scheduling of the resources of many logistics enterprises, is carried out to realize the scale and intensification of regional end distribution operations and to better adapt to the current economic development situation from the two aspects of efficiency and effectiveness so as to achieve low-cost and high-efficiency end distribution services. The joint distribution mode, which takes the logistics enterprise as the hub and connects the e-commerce platform with the logistics terminal distribution service, is the solution to the bottleneck problem encountered by the current e-commerce platform in the terminal distribution link. As the sharing mode of the third-party collection platform expands in the form of joining, it is difficult to standardize and standardize the work in this process. The

joint distribution model participates in a variety of enterprise levels, which is not conducive to the coordination between enterprises and the integration of resources. The intelligent express cabinet sharing mode takes the intelligent express cabinet as the carrier to connect the terminal logistics service and users; on the one hand, it saves the increase in operating costs caused by decentralized business outlets, and on the other hand, it uses mobile Internet technology to integrate offline physical resources and collect online service data, which not only reduces costs and increases efficiency but also simplifies the traditional terminal distribution service process and improves the quality of terminal distribution service. Let the terminal distribution service realize digitalization, informationization, and intellectualization.

When constructing the alliance of shared end distribution logistics enterprises, we have to consider the size of each enterprise, the core competitiveness, and the logistics demand of each region. Is it possible to build alliances among enterprises of different sizes? If so, what principles do you need to follow to build an alliance? In the formation of the enterprise alliance, how to protect the core competitiveness of the enterprise itself without affecting the interests of the whole enterprise alliance? In addition, what factors will hinder or promote the establishment and stability of logistics enterprise alliance? Answering the above questions is of great significance for building a shared terminal distribution platform for logistics enterprises. The main research of this paper is aimed at studying the cost-sharing ratio of each enterprise from the perspective of resource utilization efficiency, and most of the literature focuses on improving this ratio by increasing the risk factor or effort level on the basis of the Shapley value method. From the perspective of the sharing economy, this paper studies the impact of enterprises on cost allocation due to resource utilization efficiency from the point of view of the better promotion of resource integration.

## 2. Literature Review

*2.1. Research Status of Shared End Distribution.* Against the background of "Internet+logistics," sharing economy thinking is applied to the logistics industry, which continuously releases the resource dividend of integrated logistics. For the "theme of shared terminal distribution," scholars have mainly focused on the following three aspects: site selection for terminal distribution outlets, the optimization of terminal distribution path, and research on sharing terminal distribution mode. Perez-Mesa et al. [6] established a location selection model for FMCG's agricultural product distribution center with the lowest total cost. Zhang and Liu [7] proposed the method of a gray demand logistics distribution center location and model of a gray sales demand logistics distribution center location planning, which can solve the problem of gray sales demand in logistics distribution center location. Kexin et al. [8] adopted the node centrality index of complex network theory to evaluate the importance of existing terminal distribution outlets. Evari et al. [9] found that using friends in a social network to assist

in last-mile delivery greatly reduces delivery costs and total emissions while ensuring speedy and reliable delivery. In addition, Li et al. [10] analyzed the ability of logistics enterprises to meet the diversified needs of service objects through the differentiation of asset allocation. Moreover, Yu et al. [11] believed that the application of the ant colony algorithm has good feasibility for the distribution path optimization of logistics terminals. Tang et al. [12] solved the model based on multiphase particle swarm optimization algorithm (MPPSO) and Matlab. Masoud et al. [13] applied a genetic algorithm and hybrid genetic algorithm to calculate path optimization schemes and compared the results. At present, genetic algorithms are more commonly used to solve this type of problem [14–17].

Studies on the theme of the shared terminal distribution mode mainly include the following. Patterson et al. [18] established a carrier selection model, the results of which show that third-party logistics are more dependent on multimodal transport than are other terminal carriers. Lu et al. [19] applied the routing optimization algorithm to establish a multiagent simulation model to solve the routing optimization problem for electric vehicles. In the context of "Internet plus Internet of Things," Ryu et al. [20] proposed an integrated semantic service platform to support ontological models in various IoT-based service domains of a smart city. The driving factors for the formation of end-to-end joint distribution alliance are as follows. First, to meet the needs of e-commerce economic development, with the increasing maturity of e-commerce, with the upgrading of consumption, the end distribution business shows a substantial growth, and the traditional end distribution model can no longer adapt to large-scale offline distribution business [21]. The second is the need to reduce the cost structure. On the one hand, because of the high intensity of work and low wages, the mobility of end distribution employees is extremely high, and the human cost of end distribution remains high. On the other hand, the scale of operation and market limits the ability of a single enterprise to integrate resources. Third, the need to improve the efficiency of distribution. Due to the large number of service objects and the scattered geographical location, express delivery enterprises usually increase the distribution cost because of unreasonable distribution routes, high load rate, secondary distribution, and other problems [22]. Fourth, the needs of high-quality end distribution services; end distribution services gradually show a face trend, that is, with user experience as the core, personalized end distribution services are provided based on user portraits through big data, artificial intelligence, and other technologies. Fifth, to advocate the need of green logistics, the establishment of end-to-end joint distribution alliance is conducive for the integration of logistics infrastructure and other resources, and for reasonable planning of vehicle transport routes, so as to reduce traffic pollution and achieve energy saving and emission reduction.

*2.2. Research Status of the Benefit Distribution of Logistics Alliances.* Releasing the dividend of logistics integration enables end distribution enterprises to reduce costs and

increase efficiency. The design and operation of logistics alliances are critical, especially the distribution of benefits. The current research in this field is mainly classified as below.

Based on the Shapley value method and modified to put forward a benefit distribution strategy, an increasing number of scholars have used the Shapley value method to conduct benefit distribution research [23, 24]. Gong et al. [25] examine the production coordination problem from the perspective of asymmetric information. Wei [26] constructed a Shapley value model of income distribution in the information flow of a construction supply chain under asymmetric information to ensure the security of the synchronous control of information flow and a more balanced benefit distribution. Pan and Fang [27] established an integrated energy service benefit distribution model based on the improved Shapley value method to break the imbalance between the environment and economy of integrated energy services to realize overall sustainable development. Improving the Shapley value method can reduce the economic loss caused by downtime and promote multi-factory cooperation [28]. In addition, through the application of the improved Shapley value method, the production and delivery time cycle will be shortened and found that it can be applied to any industry of concern to improve operations [29].

There are many studies on the distribution of multi-agent interests, and the methods used are very diverse. Lian [30] through constructing a fuzzy time window-based scheduling optimization model for the cross-docking of cold chain logistics found that the distribution time of cold chain logistics can be significantly shortened and the distribution cost, damage cost, and economic benefits can be significantly reduced. Ma et al. [31] established a cold chain dynamic game model involving a milk manufacturer and two downstream oligopoly supermarkets under a wholesale price contract. Sun et al. [32] proposed a two-layer programming model to find the optimal location of a logistics distribution center by considering the advantages of customers and logistics planning departments. Some scholars have used the benefit distribution model of coastal port intelligent logistics supply chains based on intelligent bionic swarm evolution optimization control [33–35] or the profit distribution model based on agents [36–38] to realize the benefit distribution among multiple agents. Nadia et al. [39] through investigation and analysis proposed the necessity and importance of establishing alliances for enterprises and analyzed the interests of an alliance as a whole by using the Shapley value method. Meng et al. [40] studied multiorganization cooperation under the sharing economy model, which provided a theoretical basis for core enterprises to formulate reasonable income distribution strategies and promote the sustainable development of the sharing economy. Shang et al. [41] adopted cooperative game theory and the Shapley value method to solve the benefit distribution scheme, which improved overall risk management ability and business stability and stimulated the development potential of the enterprise.

On the whole, scholars focus on the macro research and analysis of the mode construction of shared end distribution, and pay more attention to the path optimization and site selection in shared end distribution on the micro level, but rarely analyze and study the problems in the practice of shared end distribution from the perspective of enterprise cooperation. How to complete the terminal distribution operation with the lowest cost, the least manpower, and the highest efficiency is the main problem that express delivery enterprises are facing. At present, the repeated construction of business outlets leads to the waste of social resources, the recruitment difficulties of grass-roots express salesmen, and the secondary distribution caused by the timeliness of information, making the “lone Wolf” tactic easy to suffer losses in the terminal distribution service, so it is imperative to jointly build and share logistics. It is very important to design a fair and reasonable benefit distribution mechanism to realize the sustainable development of sharing terminal distribution alliance.

### 3. Research on the Benefit Distribution Mechanism of Shared End Distribution

This paper investigates the factors that affect the benefit distribution of enterprises from the perspective of groups and individuals, as shown in Table 1, and gives the income distribution model of shared end distribution, as shown in Table 2, as follows:

When rational people participate in games involving economic activities, the two problems that need to be solved are that related to cooperation and that related to coordination. The cooperation mode is defined as the negotiating and signing of agreements to define the responsibilities and obligations of participants, and the coordination mechanism to achieve stable income expectations through good institutional design provides incentives and signals for all parties to maintain cooperation loyalty. Different cooperation modes mean that cooperative organizations have different degrees of alliance cohesion, and coordination mechanisms with different stable income expectations have different incentives and signals for members to provide cooperation. In addition, because the problems of cooperation and coordination are the two sides of game behavior, respectively, a certain cooperation model is bound to correspond to a coordination mechanism. As shown in Figure 1, the cooperation type of outsourcing business in the form of a market contract corresponds to the fixed wage model that allows for payment for a certain volume of business according to the settlement scheme; the strategic alliance model that aims to dilute the enterprise boundary to realize the sharing of management rights corresponds to the payment mode of land rent based on the distribution of alliance income on the basis of fixed payment. The ecological competition type with Internet technology as the background corresponds to the shared profit model, which is based on the coconstruction and sharing of logistics infrastructure and information resources as well as the benefits of the alliance.



TABLE 1: Influencing factors of enterprise benefit distribution in end-to-end distribution.

	Influencing factors	Meaning
Group.Visual angle	Strength and status	It determines the size of the bargaining chips of the members in the distribution of interests
	Service unit price and demand	It determines the bottom line in terms of cost that the enterprise alliance can bear and the ceiling of profit expectation
	Alliance stability	It determines whether the alliance cooperation relationship is long term or short term
Individual.Visual angle	Cost input	It determines the proportion of benefit distribution among participants
	Risk bearing	It can be divided into external and internal risk
	Service level	It determines user satisfaction
	Marginal contribution	The higher the marginal contribution rate is, the higher the sovereign status in the distribution of interests

TABLE 2: Income distribution model of shared end distribution.

	Income distribution model	Concrete
Cooperation mode of sharing resources	Outsourcing collaboration	Outsourcing of noncore logistics business
	Strategic alliance type	Dilution of enterprise boundaries through organizational cooperation agreements
	Ecological competition and cooperation	Drawing of lessons from the social organization of human beings and the operating mechanism of natural ecosystem
Coordination mechanism of income distribution	Fixed wage model	Expected coordination mechanism in the environment of outsourcing collaboration
	Mode of paying land rent	Specific monetary quantification of the benefits from resource acquisition
	Shared profit model	Emphasize that all parties should participate in the construction of logistics infrastructure and share the benefits of such cooperation

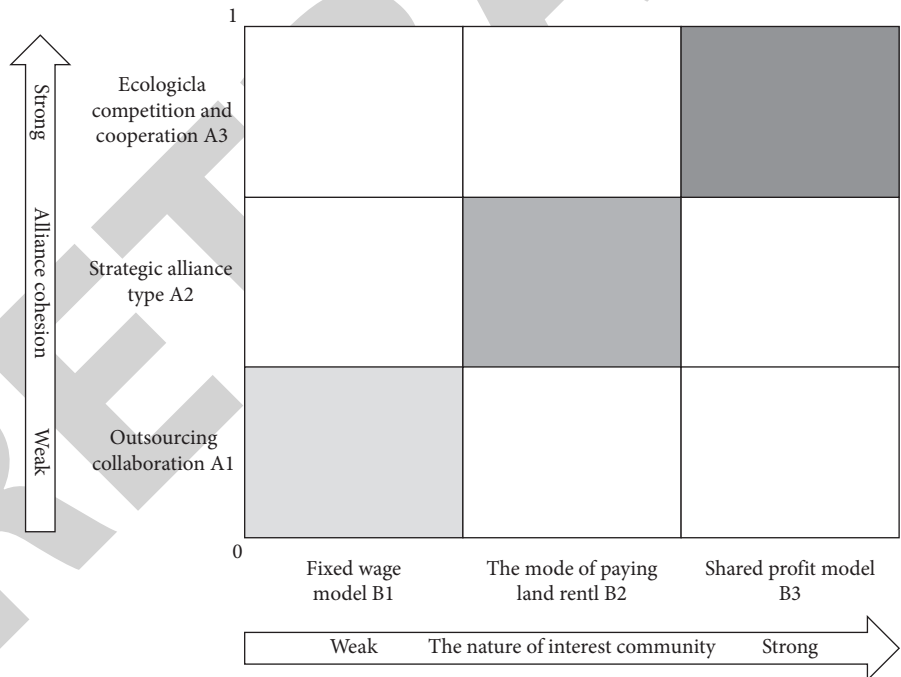


FIGURE 1: Income distribution mode selection matrix.

#### 4. The Benefit Distribution Scheme of the Modified Shapley Value Method

The Shapley value method is proposed by L. S. Shapley to solve the n-person cooperative game problem. The application of this mathematical algorithm to the benefit distribution of alliance members can effectively avoid

egalitarianism in distribution. The contribution ability of the members of the cooperative organization is included in the income distribution, which effectively arouses the production enthusiasm of the members. If  $n$  individuals participate in a social and economic activity, when the cooperative nature of the parties is greater than the competitive nature, the parties can obtain higher benefits by forming a

cooperative relationship than under the conditions of individual operation, and with the increase of the scale of the cooperative organization, the industrial synergy effect based on resource sharing and information exchange will become more prominent, and the income of the cooperative organization will gradually increase. A reasonable distribution scheme will provide corresponding organizational incentives to the members of the cooperative organization, and then ultimately ensure that the interests of all n-person cooperation can be maximized.

**4.1. Shapley Value Method Based on Contribution Degree Distribution.** For  $f$  a weighted voting scheme used by  $n$  voters to choose between two candidates, the  $n$  Shapley-Shubik Indices (or Shapley values) of  $f$  measure how much control each voter can exert over the overall outcome [42]. Suppose that there are  $n$  individuals participating in cooperation, and all participating members are denoted by  $N = \{1, 2, \dots, n\}$ ; any subset  $S$  of  $N$  (any cooperative combination in n-person cooperation) has a corresponding real-value function  $V(S)$ , that is, the benefits that can be obtained by any cooperative combination  $S$ .  $V(\emptyset) = 0$ , which means that if there are no participants in the cooperative organization, then the total income of the alliance is 0. Suppose that  $S_1$  and  $S_2$  are any two kinds of cooperative combinations in  $N$  and satisfy  $S_1 \cap S_2 = \emptyset$ . The profit created by an alliance formed by these two combinations is greater than the sum of the profits generated by these two combinations; that is,  $V(S_1 \cup S_2) > V(S_1) + V(S_2)$ . Thus,  $[N, V]$  is called the n-person cooperation strategy set. In a set of allocation strategies, it is assumed that  $X_i$  indicates that the maximum income  $V(N)$  obtained by member  $N_i$  in the alliance is a share of income, that is, the Shapley value, and the distribution strategy set of the alliance in the cooperative game is represented by  $X = \{X_1, X_2, \dots, X_n\}$ . The distribution set should meet two basic constraints—overall and individual rationality. That is, the sum of the benefits of the alliance as a whole is equal to the sum of the benefits of each member, and the income of individual members in the alliance is greater than that of individual members who complete the task alone, which is expressed by the following mathematical formula:

$$\begin{aligned} \text{overall rationality: } \sum_{i=1}^n x_i &= V(N), \\ \text{Integrated rationality: } X(i) &> V(i). \end{aligned} \quad (1)$$

According to the basic theorem of the Shapley value method, in the cooperative organization of  $n$  people, under characteristic function  $V$ , the income of each member is expressed by  $X = \{X_1, X_2, \dots, X_n\}$ . According to the Shapley value method, the income of each member can be obtained as follows:

$$X_i = \sum_{s \in S_i} W|S| [V(S) - V(S - i)]. \quad (2)$$

The formula explains the following:

- (1)  $|S|$  represents the size of the organization
- (2)  $W(|S|)$  is a weighting factor,  $W(|S|) = ((s) - 1!(n - |S|)!/n!)$ .
- (3)  $V(S - i)$  represents the benefits of the organization when member  $i$  is removed from organizational alliance  $S$
- (4)  $V(S) - V(S - i)$  represents the contribution made by member  $i$  in the organizational alliance.

According to the above introduction to the Shapley value method, we know that this method takes the contribution of each member as the distribution standard in the benefit distribution, and the premise behind this distribution method being considered reasonable and fair is that the members of the organizational alliance conform to the rational person hypothesis, and the costs and risks borne within the organizational alliance are consistent. However, compared with the specific practice, due to the differences between enterprise qualification and the competitive environment, there is a gap between the cost and risk shared by enterprises with different strengths in the organizational alliance because of the differences in their competitiveness. This kind of risk and cost gap can be shown through the cooperation mode and income distribution mechanism in the organizational alliance. When calculating members' income by the Shapley value method, it is assumed that members' ability to share the cost is consistent with the risk, which leads to the final calculation result of the Shapley value method deviating from the facts and being unable to provide the corresponding compensation mechanism for those enterprises that bear high costs and risks. Therefore, this paper hopes to modify the results calculated by the Shapley value method by measuring the evaluation indicators of different cooperation modes and income distribution mechanisms and forming weighted factors to obtain a more reasonable benefit distribution scheme that takes into account the costs and risks borne by enterprises in the cooperative organization.

**4.2. Revision of the Shapley Value Method and Establishment of the Model.** In this paper, the evaluation standard of enterprises under an established cooperation mode is based on the influencing factors of benefit distribution from the perspective of groups and evaluates the influencing factor  $A_i$  of an enterprise on organizational income output under this cooperation system from three aspects: the strength and status of alliance members, the unit price and market demand of services, and the stability of the logistics alliance relationship. The cooperation mode of sharing resources explores the influencing factors of interest distribution at the macro level of cooperative organizations, while the coordination mechanism of income distribution explores those at the micro level of alliance members. From the perspective of individual enterprises, the influencing factors of benefit distribution mainly have the following four levels: cost input, risk taking, service level, and marginal contribution. The correction results for the two angles are shown in Tables 3 and 4.

TABLE 3: Enterprise evaluation system from a group perspective.

Influencing factors of benefit distribution from a group perspective	Weight	Enterprise 1	Enterprise 2	Enterprise 3
Strength and status of the members of the alliance	0.5	$a_{11}$	$a_{21}$	$a_{31}$
Service unit price and market demand	0.3	$a_{12}$	$a_{22}$	$a_{32}$
Stability of the logistics alliance	0.2	$a_{13}$	$a_{23}$	$a_{33}$
Correction factor of enterprise i	—	$A_1$	$A_2$	$A_3$

TABLE 4: Enterprise evaluation system from an individual perspective.

Influencing factors of benefit distribution from an individual perspective	Weight	Enterprise 1	Enterprise 2	Enterprise 3
Cost input	0.3	$b_{11}$	$b_{21}$	$b_{31}$
Risk bearing	0.3	$b_{12}$	$b_{22}$	$b_{32}$
Service level	0.3	$b_{13}$	$b_{23}$	$b_{33}$
Marginal contribution	0.1	$b_{14}$	$b_{24}$	$b_{34}$
Correction factor of enterprise i	—	$B_1$	$B_2$	$B_3$

The influencing factors of the three cooperation modes on the income distribution strategy are expressed by  $A_i = \{A_1, A_2, A_3\}$ , and the following conditions are met:

- (1)  $-1 < A_i < 1; -1 < a_{ij} < 1$
- (2)  $\sum_{j=1}^3 a_{ij} = 1$
- (3)  $\sum_{j=1}^3 A_j = 0$

The influencing factors of the three income distribution coordination mechanisms on the income distribution strategy are expressed by  $B_i = \{B_1, B_2, B_3\}$ , and the following conditions are met:

- (1)  $-1 < B_i < 1; -1 < b_{ij} < 1$
- (2)  $\sum_{j=1}^3 b_{ij} = 1$
- (3)  $\sum_{j=1}^3 B_j = 0$

#### 4.3. Establishment of the Modified Shapley Value Model.

The Shapley value has become popular in the Explainable AI (XAI) literature, thanks, to a large extent, to a solid theoretical foundation, including four ‘‘favourable and fair’’ axioms for attribution in transferable utility games. The Shapley value is probably the only solution concept satisfying these axioms [43]. From the overall rational point of view, this paper probes into correction factor  $A_i$  of each enterprise in three aspects: the strength and status of alliance members, the unit price of service and market demand, and the stability of logistics alliance relationships from a group perspective. Yokote [44] introduced a new axiom, called equilibrium contribution property of equal contributors, and proved that this axiom characterizes a new class of solutions together with validity and weak covariance, providing a new axiom basis for analyzing variations of Shapley values in a uniform way. Brin [45] proved that the Shapley value is characterized by this fairness, efficiency, and the nature of the person in the empty game, and these three axioms also characterize the Shapley value of a simple game class. From the perspective of individual rationality, this paper studies the correction factor  $B_i$  of enterprises under the four dimensions of cost input, risk taking, service level, and marginal sharing. To make the final quantitative results more in line with the specific situation in practice, the

evaluation indicators of all levels and dimensions are divided into different weights, thus reflecting the unequal impact of various factors on the final benefit distribution scheme [46, 47]. After quantification, we obtain correction factor  $A_i$  of each enterprise under overall rationality and the correction factor of individual rationality  $B_i$ . By taking the average of both, we finally obtain the comprehensive correction factor of each enterprise, that is,  $1/2 * (A_i + B_i)$ . Through the comprehensive correction factor calculation, we can obtain the correction value that should be obtained by each enterprise and finally obtain the benefit distribution scheme.

In a cooperative organization of  $n$  people, under characteristic function  $V$ , the benefit of each member is expressed by  $X = \{X_1, X_2, \dots, X_n\}$ . Taking the Shapley value as the basic original value  $X_i$ , the cooperative income distribution strategy is modified from the two aspects of the cooperation mode of shared resources and the coordination mechanism of income distribution; correction factor  $1/2 * (A_i + B_i)$  is obtained; and correction value  $\Delta X_i$ ,  $\Delta X_i = (1/2) * X_i * (A_i + B_i)$  is obtained. The income of each member after the amendment is expressed by  $X' = \{X'_1, X'_2, \dots, X'_n\}$ ,  $X'_i(V) = X_i(V) + \Delta X_i(V)$ . The specific correction process is shown in Figure 2.

## 5. Example Analysis

5.1. Enterprise End-to-End Distribution Costs. Bi et al. [48] conducted an empirical study on five express delivery enterprises in China and made a comparative analysis of the results. Li [49] analyzed the problems existing in logistics distribution of express delivery enterprises, discussed various factors affecting customer satisfaction, and put forward empirical analysis assumptions. So the example analysis of this paper selects express delivery enterprises under joint distribution as the research objects and applies the correction method to modify the initial Shapley value based on the Shapley value method to obtain a more realistic distribution scheme. This paper takes 10 distribution outlets of Yuantong, Zhongtong, and Best Express enterprises in Haidian District as the research objects and studies the distribution costs borne by each enterprise under three

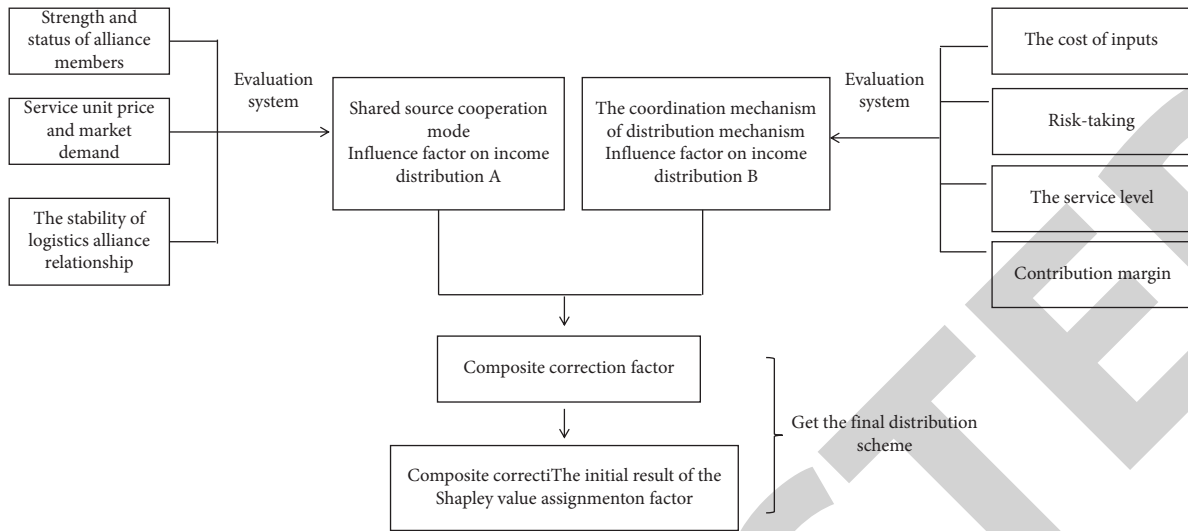


FIGURE 2: Modification process of the Shapley value method.

forms of distribution: three enterprises participate in the distribution task alone; two enterprises cooperate to participate in the joint distribution, and another enterprise participates in the distribution alone; and three enterprises cooperate to participate in the joint distribution. The corresponding distribution outlets and their numbers are shown in Table 5.

5.1.1. *Three Enterprises Participate in Distribution Separately.* The scheme and cost for the three enterprises to participate in distribution separately is shown in Table 6:

5.1.2. *Two Enterprises Cooperate to Participate in Distribution.* A. Yuantong cooperates with Zhongtong to participate in distribution, and Yunda distributes separately, the distribution path and cost are shown in Table 7:

B. Yuantong and Yunda cooperate to participate in distribution, and Zhongtong distributes separately, the distribution path and cost are shown in Table 8:

C. Yunda cooperates with Zhongtong to participate in distribution, and Yunda distributes separately, the distribution path and cost are shown in Table 9:

5.1.3. *Three Enterprises Participate in Joint Distribution.* The distribution cost under the cooperation of the three enterprises is shown in Table 10:

5.2. *Distribution of Alliance Income by the Shapley Value Method.* By calculating the costs of distribution schemes under the above cooperation modes, we can obtain the costs shared by enterprises in different distribution schemes and the benefits generated by joint distribution. On this basis, the initial Shapley values of the three enterprises after forming a cooperative organization can be calculated. To facilitate the latter description, Yuantong, Zhongtong, and Yunda enterprises are replaced by enterprises 1, 2, and 3, respectively.

As can be seen from Table 11, the participation of enterprises in joint distribution can effectively help them generate organizational cooperation income, with all three companies undertaking distribution tasks alone as the lowest threshold. When the number of enterprises participating in joint distribution continues to increase, the overall cooperation income increases significantly. In the example analysis of this paper, the total cost of the three enterprises participating in the distribution alone is 285.95. However, when the two enterprises cooperate, they obtain the benefits of the cooperative organization relative to the other enterprise that does not participate in the cooperation. Cooperation between Yuantong and Zhongtong can generate 91.1 units of income, and that between Yuantong and Yunda enterprises can generate 106.15 units of income. Moreover, cooperation between Zhongtong and Yunda enterprises can produce 89.2 units of benefits. From the perspective of bilateral cooperation, the profits of enterprises are better when they cooperate with stronger enterprises. The income from cooperation among Yuantong, Zhongtong, and Yunda is higher than that between Zhongtong and Yunda. From the perspective of multilateral cooperation, with the increase in the number of cooperative enterprises, the income of cooperative organizations also significantly increases. When the three enterprises work together to participate in a joint distribution, the terminal distribution cost can be significantly reduced by 60.6%. As a result, the profit ceiling of each enterprise at the end of distribution is raised. As large logistics enterprises gradually complete the primitive accumulation of enterprise development, gradually move toward the intelligent logistics mode in the unmanned era, and eliminate the dependence on human resources, small and medium-sized logistics enterprises, due to limited funds and a lack of technology, on the one hand, cannot realize the reform of operation mode in a short period of time. On the other hand, if enterprises continue to follow the traditional logistics management model, then they still face the constraints of human costs, so the living space of

TABLE 5: Distribution outlets and corresponding numbers.

Serial number	1	2	3	4	5
Distribution network	Yongfeng 1	China Agricultural University 2	Space bridge 3	Learn to clear the way 4	Beijing University of Aeronautics and astronautics 5
Serial number	6	7	8	9	10
Distribution network	Tsinghua university 6	Shangdi 7	Xiangshan 8	Zhongguancun 9	Longevity road 10

TABLE 6: Distribution costs of individual operations in enterprises.

Each enterprise distributes separately	Distribution route	Mileage	Traffic demand	Call vehicle	Loading rate (%)	Distribution cost
Yuantong	0-4-2-6-9-0	42.2	423.30	3	84.66	103.7
	0-7-1-0	40.5	179.31		35.86	
	0-5-3-8-10-0	64.7	441.70		88.34	
Zhongtong	0-2-1-7-0	30.3	261.25	3	52.25	86.65
	0-6-5-9-0	17.7	249.90		49.98	
	0-4-3-8-10-0	65.3	332.15		66.43	
Yunda	0-5-4-6-0	25.9	274.16	3	54.83	95.6
	0-3-2-9-7-0	46.5	325.64		65.13	
	0-10-8-1-0	58.8	260.52		52.10	

TABLE 7: Distribution costs under cooperation between Yuantong and Zhongtong

Yuantong-Zhongtong joint distribution	Distribution route	Mileage	Traffic demand	Transfer vehicle	Loading rate (%)	Distribution cost
Yuantong	0-4-5-6-9-7-0	47.5	964.38	2	96.44	43.75
Zhongtong	0-2-10-3-8-1-0	71	835.97	2	85.40	55.5
Yunda	0-5-4-6-0	25.9	274.16	3	54.83	95.6
	0-3-2-9-7-0	46.5	325.64		65.13	
	0-10-8-1-0	58.8	260.52		52.10	

TABLE 8: Distribution costs under cooperation between Yuantong and Yunda.

Yuantong-Yunda joint distribution	Distribution route	Mileage	Traffic demand	Transfer vehicle	Loading rate (%)	Distribution cost
Yuantong	0-4-2-6-5-9-0	40.5	939.94	2	93.99	40.25
Yunda	0-3-8-10-7-1-0	65.8	964.69	2	96.47	52.9
	0-2-1-7-0	30.3	281.25		56.25	
Zhongtong	0-6-5-9-0	17.7	279.90	3	55.98	86.65
	0-4-3-8-10-0	65.3	332.15		66.43	

small and medium-sized logistics enterprises is gradually narrowed. In this paper, through an analysis of examples, we know that when small and medium-sized enterprises are facing strong competitors such as large logistics enterprises, if they want to overtake them, then the establishment of a cooperative cooperation model and benefit distribution coordination mechanism suitable for the development of enterprises is necessary for their long-term development.

According to the above model of the Shapley value method and combined with the research object in the example analysis, we know that our research objects are three express delivery enterprises, namely, enterprises 1, 2, and 3, represented by the set  $N = \{1, 2, 3\}$ . The income earned by each enterprise in the cooperative organization is expressed by  $X = \{X_1, X_2, X_3\}$ . The calculation process and results are shown in Tables 12–14.

According to the above calculations, under the Shapley value model, the income of each enterprise participating in the joint distribution is as follows:

$$\begin{aligned}
 X_1 &= 15.19 + 17.7 + 28.98 = 61.86, \\
 X_2 &= 15.19 + 14.87 + 23.33 = 53.39, \\
 X_3 &= 17.7 + 14.87 + 28.35 = 60.91.
 \end{aligned}
 \tag{3}$$

In the distribution scheme of the Shapley value model, there is little difference in the organizational income of the enterprises participating in the joint distribution, which stems from the assumption at the beginning of the establishment of the model: the differences among participants should be ignored. However, due to the different qualifications of each enterprise, the contribution ability of the heterogeneous enterprise resources to the cooperative organization is different. When the heterogeneous enterprise resources help the cooperative organization achieve high returns, this part of the income should be compensated to the enterprise in terms of contribution ability to provide the cooperative organization with the development incentive of “contributing enterprise resources.” Only in this way can the sustainable and coordinated development of the cooperative organization be realized.

TABLE 9: Distribution costs under cooperation between Yunda and Zhongtong

Zhongtong-Yunda joint distribution	Distribution route	Mileage	Traffic demand	Transfer vehicle	Loading rate (%)	Distribution cost
Yunda	0-5-4-6-2-9-0	37.4	895.94	2	89.59	38.7
Zhongtong	0-7-1-3-8-10-0	68.7	924.69	2	92.47	54.35
	0-4-2-6-9-0	42.2	423.30	3	84.66	103.7
Yuantong	0-7-1-0	40.5	179.31		35.86	
	0-5-3-8-10-0	64.7	441.70		88.34	

TABLE 10: Distribution costs under the cooperation of three enterprises.

Joint distribution by three enterprises	Distribution route	Mileage	Traffic demand	Transfer vehicle	Loading rate (%)	Distribution cost
Yuantong	0-4-5-6-0	33.9	912.48	1	91.25	36.95
Zhongtong	0-2-1-7-9-0	29.9	998.35	1	99.83	34.95
Yunda	0-3-8-10-0	35.8	954.11	1	95.41	37.9

TABLE 11: Organizational benefits under various modes of cooperation.

Distribution plan	Total distribution cost	Revenue from joint distribution
Three separate distributions	285.95	0
Joint distribution between enterprises 1 and 2	194.85	91.1
Joint distribution between enterprises 1 and 3	179.8	106.15
Joint distribution between enterprises 2 and 3	196.75	89.2
Three joint distributions	109.8	176.15

TABLE 12: Calculation table of the initial Shapley value of the Yuantong enterprise.

S	{1}	{1, 2}	{1, 3}	{1, 2, 3}
$V(S)$	0	91.1	106.15	176.15
$V(S - \{1\})$	0	0	0	89.2
$V(S) - V(S - \{1\})$	0	91.1	106.15	86.95
$ S $	1	2	2	3
$\omega(S)$	1/3	1/6	1/6	1/3
$\omega(S)[V(S) - V(S - \{i\})]$	0.00	15.19	017.70	29.98

5.3. *Modification of the Initial Shapley Value.* The following is the correction of the initial Shapley value of each enterprise obtained through the example analysis. The source of the correction factor is based on the evaluation indicators of each enterprise in three aspects: the strength and status of alliance members, service unit price and market demand, and the stability of the logistics alliance relationship. The feedback source of the evaluation index of the enterprise is mainly obtained by network enterprise data and offline interviews and quantifies the evaluation index of each enterprise according to first-hand data.

The initial Shapley value of the enterprise is modified from the point of view of the cooperation mode of sharing resources, and the details of the correction factors of each enterprise are shown in Table 15.

The initial Shapley value of the enterprise is modified from the point of view of the coordination mechanism of income distribution, and the correction factors of each enterprise are shown in Table 16.

Through the above calculations, the final benefit distribution correction factors of the three enterprises participating in joint distribution can be obtained. To make the correction results more reasonable by averaging the

correction factors from two perspectives, by calculating the final correction factors of each enterprise, correction value  $\Delta Xi, \Delta Xi = Xi * 1/2 (Aj + Bj)$  is finally obtained. The specific results are shown in Table 17.

Yuantong Express is ahead of the other two enterprises in terms of overall market share and service quality, and its business orders are also greater in amount and number than are those of the other two enterprises, which will make an important contribution to the income of the whole cooperative organization in the process of building the alliance. Yunda Express has weak market share and resource scheduling ability compared with Yuantong and Zhongtong enterprises, showing positive externalities to organizational income. With the help of cooperative organizations, the Yunda enterprise reduces the cost of terminal distribution and obtains the opportunity to share the benefits from cooperation, but to compensate for the contribution of Yuantong and Zhongtong, part of its income needs to be allocated to them to send a positive incentive signal to them in terms of contribution. Therefore, by observing the correction factors of various enterprises, we can see that those of the Yuantong and Zhongtong enterprises are positive, while the correction factor of the Yunda enterprise is negative.



TABLE 13: Calculation table of the initial Shapley value of the Zhongtong enterprise.

S	{2}	{2, 1}	{2, 3}	{1, 2, 3}
$V(S)$	0	91.1	89.2	176.15
$V(S - \{2\})$	0	0	0	106.15
$V(S) - V(S - \{2\})$	0	91.1	89.2	70
$ S $	1	2	2	3
$\omega(S)$	1/3	1/6	1/6	1/3
$\omega(S)[V(S) - V(S - \{i\})]$	0.00	15.19	14.87	23.33

TABLE 14: Calculation table of the initial Shapley value of the Yunda enterprise.

S	{3}	{3, 1}	{3, 2}	{1, 2, 3}
$V(S)$	0	106.15	89.2	176.15
$V(S - \{3\})$	0	0	0	91.1
$V(S) - V(S - \{3\})$	0	106.15	89.2	85.05
$ S $	1	2	2	3
$\omega(S)$	1/3	1/6	1/6	1/3
$\omega(S)[V(S) - V(S - \{i\})]$	0.00	17.70	14.87	28.35

TABLE 15: Correction factors of enterprises from the perspective of groups.

Cooperation mode of shared resources	Weight	Enterprise 1	Enterprise 2	Enterprise 3
Strength and status of the members of the alliance	0.5	0.53	0.32	-0.85
Service unit price and market demand	0.3	0.83	0.76	-1.59
Stability of logistics alliance	0.2	-0.43	0.26	0.17
Correction factor of enterprise i	—	0.43	0.44	-0.87

TABLE 16: Correction factors of each enterprise from an individual perspective.

Coordination mechanism of income distribution	Weight	Enterprise 1	Enterprise 2	Enterprise 3
Cost input	0.3	0.54	0.34	-0.88
Risk bearing	0.3	0.36	0.27	-0.63
Service level	0.3	0.47	0.37	-0.84
Marginal contribution	0.1	0.31	0.29	-0.6
Correction factor of enterprise i	—	0.44	0.32	-0.77

TABLE 17: Corrections for each enterprise.

	Yuantong	Zhongtong	Yunda
Correction factor	0.31	0.24	-0.55
Correction value	18.90	12.95	-33.38

TABLE 18: Comparison of results before and after correction.

	Yuantong	Zhongtong	Yunda
Shapley value method	61.86	53.39	60.91
After correction	18.90	12.95	-33.38

Since the main purpose of this step is to calculate the correction value of each enterprise, the correction value further modifies the distribution scheme on the basis of that obtained by the Shapley value method, so the sum of the correction value of each enterprise should be equal to 0; that is, the correction scheme of the paper does not change the income of the cooperative organization in the whole correction process. On the basis of this modified value, we can obtain the final distribution scheme by adding the initial Shapley value calculated by the Shapley value method, and the final distribution scheme is compared with that obtained by the Shapley method, as shown in Table 18:

Through the above comparison, we can see that the distribution scheme of the Shapley value method pays more attention to the distribution of cooperative income from the perspective of average distribution, and there is little difference in the respective income of the three enterprises. This "average" distribution method does not give

compensation and incentives to the contribution ability of enterprises with different qualifications, which is not conducive for the sustainable development of cooperative organizations. In the example analysis of this paper, we can see that based on the distribution scheme according to the Shapley value method, the cooperative benefits shared by the three enterprises are not very different, and the difference is controlled within 9 units. The distribution scheme under the average principle cannot reflect the utility of each individual in the cooperative organization, quantify this utility, and reward this utility through reasonable benefit compensation. In the example analysis of this paper, the Yuantong enterprise can bring higher business order and resource scheduling ability to the cooperative organization and can share its offline logistics resources with other enterprises; after revision, the income of the Yuantong enterprise is obviously higher than that of the Zhongtong and Yunda enterprises.

The case study of this paper takes joint distribution as the research background, so in general enterprise practice, we usually choose the cooperation mode of diluting enterprise boundary, which is higher than outsourcing cooperation. The core of collaborative joint distribution is sharing orders. When strong enterprises share internal low-profit orders with members of cooperative organizations, while other members complete these orders to harvest enterprise profits, it is necessary to pay a corresponding proportion of “rent” to stronger enterprises, which is also the basic idea of the mode of paying land rent in the coordination mechanism of income distribution. This kind of “rent” incentive is to make stronger enterprises more willing to participate in the construction of cooperative organizations, so as to maintain the sustainable development of cooperative organizations. In this case, because of its limited market share and service level, Yunda enterprise undertakes the terminal distribution business of high cost and low volume when it operates alone. However, when the distribution business is optimized through joint distribution, from its own vertical comparison, Yunda enterprise not only completes the corresponding distribution tasks but also obtains the benefits of cooperation. However, the stability of this form of cooperation depends on the comparison of the strength of their respective enterprises, which is mainly reflected in the market share; comprehensive service capacity; and complete infrastructure. In terms of logistics resource scheduling ability, when the strength of enterprises is very different and the balance of cooperation is broken, enterprises with greater discourse power will force weaker enterprises to accept organizational agreements that are not conducive for the development of enterprises, which is likely to break the calm mode of cooperation.

## 6. Conclusions

This paper focuses on the difficult problem of enterprise cooperative benefit distribution in end distribution, takes the benefit distribution principle as the basic constraint, and constructs a two-dimensional enterprise evaluation system from the group and individual perspectives. The income distribution mode selection matrix is constructed from the two dimensions of the cooperation mode and coordination mechanism. Under these different practical backgrounds, all enterprise parties choose the appropriate cooperation mode and income distribution mode in the income distribution mode selection matrix and obtain the income distribution correction factor from the evaluation system within the selected range, which is modified on the basis of the initial Shapley value. Compared with the initial Shapley value method, which pursues income equalization, the revised income distribution results more so reflect the income-generating ability of different enterprises in the cooperative organization and give different proportions of benefits to the cooperative organization based on different income-generating capacities to provide incentives for the development of the cooperative organization. This paper holds that the influence of the income distribution scheme of cooperative organizations depends on the close relationship between the

cooperation mode of shared resources and the coordination mechanism of income distribution. Under this two-dimensional mechanism, the final benefit distribution should be determined based on the evaluation of the enterprise and not on only the Shapley value mathematical model. The two-dimensional selection mechanism based on income distribution can better reflect the income-generating ability that enterprises can exert in the specific cooperation practice. From a rational point of view, this paper proposes what kind of cooperation mode and coordination mechanism enterprises should be established to realize the sustainable development of cooperative organizations. However, in practice, it is still necessary to make prudent cooperation decisions based on an enterprise’s internal resources, enterprise strategy, enterprise organizational structure, and the external environment.

This paper focuses on studying and solving the problem of the benefit distribution of various enterprises, puts forward a framework solution from the perspective of static research, and presents a two-dimensional selection matrix of income distribution based on overall and individual rationality. The Shapley value algorithm is modified from the point of view of the cooperation mode and income distribution coordination mechanism. Although this paper takes into account the influencing factors of overall rationality on income distribution and individual rationality on income distribution, it does not specifically analyze the influence of the negotiation premium power between different enterprises on the whole income distribution, that is, taking into account the rationality of the establishment of cooperative organizations and the rationality of individual participation in cooperative organizations, but there is no dynamic research perspective for analyzing the continuous growth of enterprises in cooperative organizations. Whether the old cooperation model and distribution mechanism can still be established needs to be further studied. Therefore, this paper does not consider the dynamic cooperative development of enterprises, and the static research perspective on the distribution of interests of enterprises inevitably makes the final research results deviate from the practices of specific enterprises.

## Data Availability

The data generated and analyzed during this research are available from the corresponding author on request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References

- [1] Y. Song, F. R. Yu, L. Zhou, X. Yang, and Z. He, “Applications of the internet of Things (IoT) in smart logistics: a comprehensive survey,” *IEEE Internet of Things Journal*, vol. 8, no. 6, pp. 4250–4274, 2021.
- [2] M. Saini and D. Hruševská, “Influence of logistics competitiveness and logistics cost on economic development: an

- FsQCA qualitative approach,” *E+M Ekonomie a Management*, vol. 24, no. 2, pp. 51–64, 2021.
- [3] I. Davydenko, L. Tavasszy, and H. Quak, “On the cost elasticity of inter-regional distribution structures: a case study for The Netherlands,” *European Journal of Transport and Infrastructure Research*, vol. 21, no. 3, pp. 1–18, 2021.
  - [4] M. J. Baker and K. J. Swope, “Sharing, gift-giving, and optimal resource use in hunter-gatherer society,” *Economics of Governance*, vol. 22, no. 2, pp. 119–138, 2021.
  - [5] D. Gong, M. Tang, S. Liu, G. Xue, and L. Wang, “Achieving sustainable transport through resource scheduling: a case study for electric vehicle charging stations,” *Advances in Production Engineering & Management*, vol. 14, no. 1, pp. 65–79, 2019.
  - [6] J. C. Pérez-Mesa, M. M. Serrano-Arcos, J. F. Jiménez-Guerrero, and R. Sánchez-Fernández, “Addressing the location problem of a perishables redistribution center in the middle of Europe,” *Foods*, vol. 10, no. 5, p. 1091, 2021.
  - [7] Q. S. Zhang and H. Liu, “Location of logistics distribution centers with Grey demand based on hybrid PSO,” *Journal of Grey System*, vol. 23, no. 3, pp. 291–298, 2011.
  - [8] B. Kexin, Y. Mengke, Z. Latif, and Z. Xiaoguang, “A new solution for city distribution to achieve environmental benefits within the trend of green logistics: a case study in China,” *Sustainability*, vol. 12, no. 20, p. 8312, 2020.
  - [9] A. Devari, A. G. Nikolaev, and Q. He, “Crowdsourcing the last mile delivery of online orders by exploiting the social networks of retail store customers,” *Transportation Research Part E: Logistics and Transportation Review*, vol. 105, pp. 105–122, 2017.
  - [10] L. I. Guoqi, F. Jin, Y. Chen, J. Jiao, and S. Liu, “Location characteristics and differentiation mechanism of logistics nodes and logistics enterprises based on points of interest (POI): a case study of Beijing,” *Journal of Geographical Sciences*, vol. 27, no. 7, pp. 879–896, 2017.
  - [11] M. Yu, G. Yue, Z. Lu, and X. Pang, “Logistics terminal distribution mode and path optimization based on ant colony algorithm,” *Wireless Personal Communications*, vol. 102, no. 4, pp. 2969–2985, 2018.
  - [12] M. Tang, D. Gong, S. Liu, and H. Zhang, “Applying multi-phase particle swarm optimization to solve bulk cargo port scheduling problem,” *Advances in Production Engineering & Management*, vol. 11, no. 4, pp. 299–310, 2016.
  - [13] M. Rabbani, H. Farrokhi-Asl, and B. Asgarian, “Solving a bi-objective location routing problem by a NSGA-II combined with clustering approach: application in waste collection problem,” *Journal of Industrial Engineering International*, vol. 13, no. 1, pp. 13–27, 2017.
  - [14] Y. Sun, N. Geng, S. Gong, and Y. Yang, “Research on improved genetic algorithm in path optimization of aviation logistics distribution center,” *Journal of Intelligent and Fuzzy Systems*, vol. 38, no. 1, pp. 29–37, 2019.
  - [15] D. E. Gomes, M. I. D. Iglésias, A. P. Proença, T. M. Lima, and P. D. Gaspar, “Applying a genetic algorithm to a m-TSP: case study of a decision support system for optimizing a beverage logistics vehicles routing problem,” *Electronics*, vol. 10, no. 18, p. 2298, 2021.
  - [16] C. Chen and S. Wu, “Multi-objective distribution routing optimization with time window based on improved genetic algorithm,” *Latin American Applied Research*, vol. 48, no. 3, pp. 151–156, 2018.
  - [17] K. Guo, “Research on location selection model of distribution network with constrained line constraints based on genetic algorithm,” *Neural Computing & Applications*, vol. 32, no. 6, pp. 1679–1689, 2020.
  - [18] Z. Patterson, G. O. Ewing, and M. Haider, “How different is carrier choice for third party logistics companies?” *Transportation Research Part E: Logistics and Transportation Review*, vol. 46, no. 5, pp. 764–774, 2010.
  - [19] X. C. Lu, Q. B. Chen, and Z. J. Zhang, “The electric vehicle routing optimizing algorithm and the charging stations’ layout analysis in beijing,” *International Journal of Simulation Modelling*, vol. 13, no. 1, pp. 116–127, 2014.
  - [20] M. Ryu, J. Kim, and J. Yun, “Integrated semantics service platform for the internet of Things: a case study of a smart office,” *Sensors*, vol. 15, no. 1, pp. 2137–2160, 2015.
  - [21] S. Hong, R. Lv, and P. Hong, “Cost sharing of terminal joint distribution of express industry,” *IET Intelligent Transport Systems*, vol. 12, no. 7, pp. 730–734, 2018.
  - [22] A. Milan, B. Neboja, and M. Kilibarda, “Benchmarking distribution centres using principal component analysis and data envelopment analysis: a case study of Serbia,” *Expert Systems with Applications*, vol. 40, no. 10, pp. 3926–3933, 2013.
  - [23] X. Liu, X. Wang, H. Guo, and X. An, “Benefit allocation in shared water-saving management contract projects based on modified expected shapley value,” *Water Resources Management*, vol. 35, no. 1, pp. 1–24, 2021.
  - [24] J. Quigley and L. Walls, “Trading reliability targets within a supply chain using Shapley’s value,” *Reliability Engineering & System Safety*, vol. 92, no. 10, pp. 1448–1457, 2007.
  - [25] D. Gong, M. Tang, S. Liu, and Q. Li, “A reconsidering production coordination: a principal-agent theory-based analysis,” *Advances in Production Engineering & Management*, vol. 12, no. 1, pp. 51–61, 2017.
  - [26] H. Wei, “Synchronization control method of information flow in coastal logistics supply chain under asymmetric information,” *Journal of Coastal Research*, vol. 93, no. sp1, pp. 1087–1092, 2019.
  - [27] H. Pan and J. Fang, “Research on benefit distribution of comprehensive energy services under the background of green development concept,” *Fresenius Environmental Bulletin*, vol. 30, no. 6, pp. 6188–6197, 2021.
  - [28] Y. Tian and S. Li, “Multi-plant direct heat integration based on the risk-based Shapley value and Alopex-based evolutionary algorithm,” *Chemical Engineering Science*, vol. 248, 2022.
  - [29] N. Kakoty, P. Baruah, and S. Borkotokey, “The role of the non-productive players in cooperative games with transferable utilities: a survey,” *International Journal of General Systems*, vol. 50, no. 5, pp. 527–547, 2021.
  - [30] J. Lian, “An optimization model of cross-docking scheduling of cold chain logistics based on fuzzy time window,” *Journal of Intelligent and Fuzzy Systems*, vol. 41, no. 1, pp. 1901–1915, 2021.
  - [31] M. Junhai, X. Tiantong, H. Yalan, and Z. Xueli, “Impact research on a nonlinear cold chain evolutionary game under three various contracts,” *International Journal of Bifurcation and Chaos*, vol. 29, no. 5, Article ID 1950058, 2019.
  - [32] H. Sun, Z. Gao, and J. Wu, “A bi-level programming model and solution algorithm for the location of logistics distribution centers,” *Applied Mathematical Modelling*, vol. 32, no. 4, pp. 610–616, 2008.
  - [33] S. Li, Q. Sun, and W. Wu, “Benefit distribution method of coastal port intelligent logistics supply chain under cloud computing,” *Journal of Coastal Research*, vol. 93, no. sp1, pp. 1041–1046, 2019.
  - [34] F. Lu, L. Wang, H. Bi, Z. Du, and S. Wang, “An improved revenue distribution model for logistics service supply chain considering fairness preference,” *Sustainability*, vol. 13, no. 12, p. 6711, 2021.

## Research Article

# Image Encryption Based on Hopfield Neural Network and Bidirectional Flipping

Haitao Zhang  and Shuangqi Yang

*School of Software, Liaoning Technical University, Huludao 125105, China*

Correspondence should be addressed to Haitao Zhang; [dalianjh@djtu.edu.cn](mailto:dalianjh@djtu.edu.cn)

Received 6 December 2021; Accepted 7 January 2022; Published 11 February 2022

Academic Editor: Daqing Gong

Copyright © 2022 Haitao Zhang and Shuangqi Yang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Many encryption systems face two problems: the key has nothing to do with the plaintext; only a single chaotic sequence is adopted during the encryption. To solve the problems, this paper proposes an image encryption method based on Hopfield neural network and bidirectional flipping. Firstly, the plaintext image was segmented into blocks, the resulting image matrix was block scrambled, and each block was bidirectionally flipped to complete the scrambling process. After that, the plaintext image was processed by the hash algorithm to obtain the initial values and control parameters of the chaotic system, producing a pseudo-random sequence. Then, a diffusion matrix was generated through the optimization by Hopfield neural network and used to derive a ciphertext image through diffusion transformation. Experimental results show that our algorithm is highly sensitive to plaintext, strongly resistant to common attacks, and very efficient in encryption.

## 1. Introduction

The safety of digital images, an important carrier of information, has attracted much interest and concern [1]. To ensure the safety of image information, it is highly necessary to develop a good encryption algorithm [2]. The chaotic system lays a good foundation for encryption systems, due to its excellent sensitivity to initial values. Due to the sensitivity of the initial value, small changes can get completely different results, so each small change can achieve completely different encryption results during encryption, which can provide good security. As a result, the chaotic system is being integrated to more and more encryption algorithms. In the past two decades, researchers have proposed various encryption methods and applied them to image encryption, drawing on the unique properties of the chaotic system (e.g., sensitivity to initial values, unpredictability, and pseudo-randomness) and the natural bound between the system and cryptology [3–10]. Some scholars put forward several new chaotic systems and designed the corresponding encryption strategies [4–8]. Some scholars presented encryption algorithms based on existing chaotic systems, focusing on the

design of encryption strategies [9–16]. Some scholars combined spatiotemporal chaos with DNA sequencing [6] and proposed image encryption algorithms based on the cryptological features of spatiotemporal nonadjacent coupled map lattices [17] and mixed linear-nonlinear coupled map lattices [18], respectively. The safety performance of an encryption algorithm can be measured by an important criterion: the ability to resist various attacks, namely, violent attack, statistical attack, and differential attack. Some algorithms are unable to withstand chosen-plaintext attack [19–24]. Besides, many algorithms are inefficient in encryption and need multiple encryptions to achieve a good effect. To solve the above defects, this paper explores key generation and image encryption strategy and proposes a chaotic image encryption algorithm based on Hopfield neural network and the image scrambling approach of bidirectional flipping. Firstly, the plaintext image was segmented into multiple  $N \times N$  blocks, and the resulting image matrix was block scrambled. Each block was bidirectionally flipped and merged into a scrambled image. Next, the plaintext image was processed by the hash function, producing a hash array. On this basis, the control parameters

and initial values of the chaotic system were determined to generate a random pseudo-matrix. Multiple sequences were taken as the initial conditions of Hopfield chaotic neural network, which creates the key flow of the diffusion matrix. Then, the scrambled image was segmented along the diagonal, and the key flow was converted into a key matrix. Afterwards, symmetric diffusion was performed on the key matrix and the scrambled image to obtain the encrypted image. Finally, the safety and reliability of our algorithm were demonstrated by comparing it with similar algorithms developed since 2017.

## 2. Image Encryption Strategy

*2.1. Hopfield Neural Network.* Proposed by American physicist Hopfield in 1982, the Hopfield neural network mimics the memory mechanism of biological neural networks. In this fully connected neural network, every node transmits a signal to other nodes, which eventually return the signal to the transmitter. Therefore, the Hopfield neural network has a feedback mechanism. A typical Hopfield neural network can be expressed as

$$x = -x_i + \sum_{i=1}^3 w_{ij}v_i, \quad (1)$$

where  $v$  is the hyperbolic tangent function:

$$v_i = \tanh(x_i) = \frac{e^{x_i} - e^{-x_i}}{e^{x_i} + e^{-x_i}}, \quad (2)$$

and  $w$  is the weight function:

$$w = \begin{bmatrix} 2 & -1 & 0 \\ 1.7 & 1.71 & 1.1 \\ -2.5 & -2.9 & 0.56 \end{bmatrix}. \quad (3)$$

*2.2. Encryption Flow.* Our algorithm calls logistic mapping repeatedly:

$$x_{n+1} = rx_n(1 - x_n), \quad (4)$$

where  $r \in (0, 4]$  is a control parameter. If  $3.5699456 \leq r \leq 4$ , the logistic mapping will be chaotic; as  $r$  gradually approaches 4, the mapping becomes more and more chaotic and generates a chaotic sequence  $x_n$  of a better quality.

Taking an  $M \times M$  plaintext image  $P$  for example, this paper designs a novel image encryption algorithm. There are three steps of the algorithm: image segmentation, block scrambling, and symmetric diffusion. After scrambling and diffusion, the plaintext image  $P$  is improved into a highly secure encrypted image.

## 3. Encryption Algorithm

Like most encryption strategies, our encryption strategy consists of two steps: scrambling and diffusion.

*3.1. Bidirectional Flipping.* Image scrambling aims to change the position of image pixels. The specific process of scrambling through block-based triangular transform is as follows.

Firstly, the original image is segmented into  $N \times N$  blocks, each of which is subjected to triangular transform. The segmented image matrix consumes less resources in the process of computing encryption and can be further scrambled. Suppose the original image is of the size  $256 \times 256$  and is broken down into  $8 \times 8$  blocks. Then, the scrambling can be realized in the following steps:

*Step 1.* The hash function is applied on the plaintext image to generate a hash array. The relevant values are extracted from the array for initialization, producing the initial values and control parameters of the chaotic system. Then, a pseudo-random sequence is generated through the logistic chaotic system and is taken as the initial values of the Hopfield chaotic neural network. The optimal random sequence is thereby obtained.

*Step 2.* After segmentation, the image matrix is scrambled with the random sequence obtained by Hopfield chaotic neural network.

*Step 3.* The random sequence is numerically calculated. The rotation direction and angle of the blocks on the first layer of the image matrix are solved through remainder operation, with  $90^\circ$  as a unit. On this basis, the scrambling model is determined for the entire matrix. The calculation formulas are as follows:

$$\begin{aligned} Z(i) &= \text{floor}[(z(i) \times 10^n) \bmod 256], \\ H(i) &= Z(i) \bmod M. \end{aligned} \quad (5)$$

*Step 4.* Each image block is transformed into a matrix. The matrix of each block is rotated clockwise or counter-clockwise. As required by the algorithm, the adjacent layers are rotated in opposite directions.

*Step 5.* The scrambled block matrices are merged to a matrix as large as the original image matrix. This is the final result of image scrambling.

*3.2. Image Diffusion.* Image diffusion mainly segments the original image matrix along the diagonal. The specific diffusion process is as follows:

*Step 1.* The initial state  $x_2$  and control parameter  $r_2$  of the logistic mapping are calculated as described in Section 2.1.

*Step 2.* Logistic mapping is implemented iteratively 200 times, producing chaotic sequences A1–A3. These sequences are imported to Hopfield chaotic neural network as initial parameters and converted into a diffusion key flow, forming a key matrix.

*Step 3.* The key matrix and scrambled image CC1 are segmented along the same direction. Suppose the matrix is of the size  $8 \times 8$ . The diagonal segmentation model is illustrated in Figure 1.

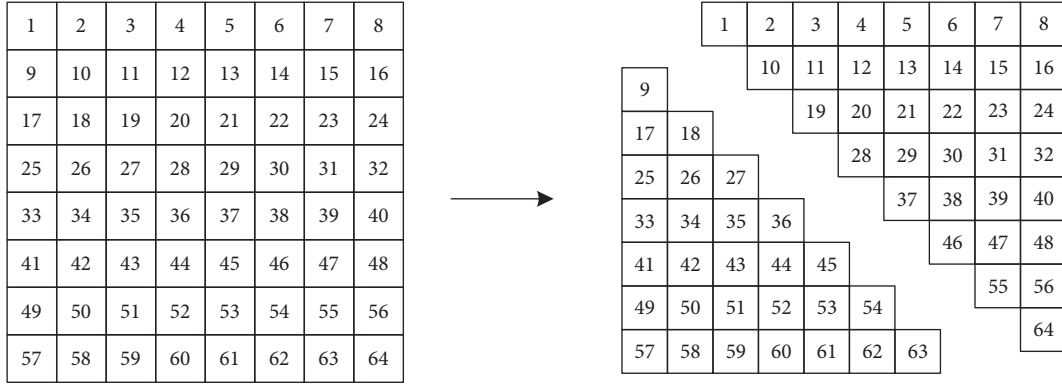


FIGURE 1: Diagonal segmentation model.

*Step 4.* After the diagonal segmentation model is determined, the diffusion is performed by (see Figure 2)

$$\begin{cases} W_T'(i, j) = W_T(i, j) \oplus Q_K(i, j) \\ B_T'(i, j) = B_T(i, j) \oplus W_T'(i, j) \end{cases} \quad (6)$$

*Step 5.* The diffusion image is obtained through the diffusion operation, and the final encrypted image  $C$  is outputted.

The decryption is the inverse process of encryption.

#### 4. Simulation Results

To verify its effectiveness, our encryption algorithm was simulated on multiple images, using the simulation software GNU Octave. Through an experiment, the initial values of logistic mapping were determined as  $x_1(0) = 0.8761$  and  $x_2(0) = 0.7323$ ; the control parameters of logistic mapping were finalized as  $r_1 = 3.9695$  and  $r_2 = 3.8925$ ; the total number of iterations (TNI) was set to 200; different hash arrays  $H$  were generated from different plaintext images.

Our simulation uses the gray image of Lena ( $256 \times 256$ ) and the color image of peppers. For the color image, firstly, the gray level of the color image is transformed, and the layer is divided into three different gray levels: R, G, and B, which are encrypted in the corresponding encryption process. Figure 3 shows the generated encrypted images and decrypted images.

#### 5. Safety Analysis

An ideal encryption algorithm should be able to resist various attacks, such as violent attack, statistical attack, and

differential attack, and chosen-plaintext attack. To verify the safety of our algorithm, this paper theoretically analyzes and numerically simulates the algorithm in five aspects and compares it with the state-of-the-art chaotic theory-based algorithms [25–30].

**5.1. Histogram Analysis.** The ability of an encryption algorithm to resist statistical attack can be directly measured by the histogram of the ciphertext image, which describes the pixel distribution of the image. The statistical attack can easily steal some information from an image with uneven pixel distribution. Image itself is a form of data information, and image itself can be used as a carrier or directly as a kind of information transmission. The pixel distribution of the original image is distributed according to the content level of the image, so important content can be stolen and obstructed through statistical analysis attacks.

It can be known that the pixels were not evenly distributed on the plaintext histograms but evenly distributed on the ciphertext histograms. Therefore, the ciphertext images obtained by our encryption algorithm can resist the statistical attack.

**5.2. Correlation Analysis.** A high correlation between adjacent pixels indicates that the plaintext image is prone to the statistical attack. Thus, it is necessary for the encryption algorithm to reduce the correlation between adjacent pixels. 10,000 pixels were randomly selected from the plaintext and ciphertext images of the gray image of Lena, respectively. Then, the correlations between adjacent pixels in the horizontal, vertical, and diagonal directions were calculated by

$$\begin{aligned} r_{xy} &= \frac{\text{cov}(x, y)}{\sqrt{D(x)}\sqrt{D(y)}}, \quad \text{cov}(x, y) = \frac{1}{N} \sum_{i=1}^N (x_i - E(x))(y_i - E(y)), \\ D(x) &= \frac{1}{N} \sum_{i=1}^N (x_i - E(x))^2, \quad E(x) = \frac{1}{N} \sum_{i=1}^N x_i. \end{aligned} \quad (7)$$



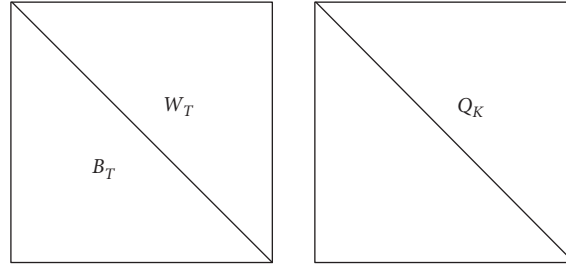


FIGURE 2: Schematic diagram of symmetrical segmentation.

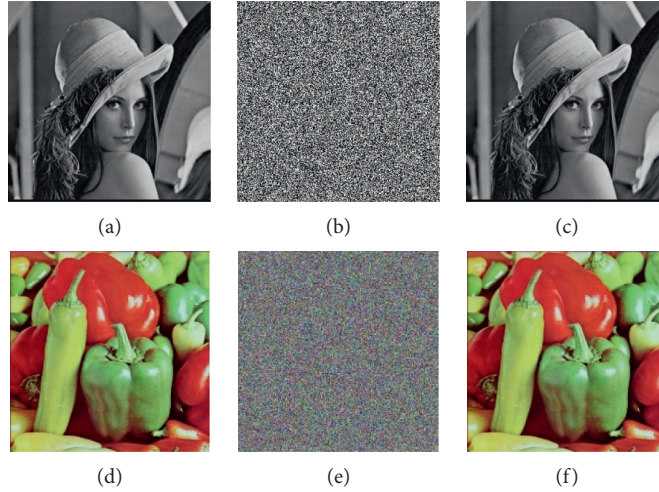


FIGURE 3: Simulation results. (a) Original image of Lena. (b) Encrypted image of Lena. (c) Decrypted image of Lena. (d) Original image of peppers. (e) Encrypted image of peppers. (f) Decrypted image of peppers.

In addition, the correlations between adjacent pixels in the ciphertext image of Lena obtained by our algorithm were compared with those in the ciphertext image of Lena obtained by other algorithms (Table 1).

**5.3. Information Entropy Analysis.** Information entropy is an important indicator of the randomness of information:

$$H(s) = \sum_{i=0}^{2^n-1} p(m_i) \log_2 \frac{1}{p(m_i)}, \quad (8)$$

where  $p(s_i)$  is the probability of  $s_i$ .

In theory, the probability of information leak decreases as the information entropy approaches 8. Table 2 compares the information entropy of the cyphertexts of two test images obtained by our algorithm with that obtained by three other algorithms [25–27]. On both test images, the information entropy was approximately 8 in the ciphertexts obtained by our algorithm. The information entropy obtained by our algorithm was closer to 8 than that of any other algorithm. Therefore, the ciphertext images obtained by our algorithm are unlikely to suffer from information leak and are robust against the statistical attack.

**5.4. Differential Attack.** Differential attack is a kind of chosen-plaintext attack. During the attack, the attacker makes minor modifications to the plaintext image, encrypts the modified image and the original image separately, and compares the two encrypted images to find the correlations between plaintext and ciphertext images. The differential attack is commonly evaluated by the number of pixel change rate (NPCR) and the unified average changing intensity (UACI):

$$\text{NPCR} = \frac{1}{W \times H} \sum_{i=1}^W \sum_{j=1}^H D(i, j) \times 100\%, \quad (9)$$

$$\text{UACI} = \frac{1}{W \times H} \left( \sum_{i=1}^W \sum_{j=1}^H \frac{|c_1(i, j) - c_2(i, j)|}{255} \right) \times 100\%,$$

where  $W$  and  $H$  are image width and height, respectively;  $c_1$  is the original plaintext image; and  $c_2$  is the plaintext image derived from  $c_1$  by changing 1 bit of pixel value. If  $c_1(i, j) \neq c_2(i, j)$ , then  $D(i, j) = 1$ ; otherwise,  $D(i, j) = 0$ .

Theoretically, the result is good if NPCR and UACI approach 99.6093% and 33.4635%, respectively. Without changing the keys, our encryption algorithm was adopted to encrypt  $c_1$  and  $c_2$ , respectively. Next, the NPCR and UACI were calculated for the two resulting ciphertext images.

TABLE 1: Correlations between adjacent pixels in ciphertext image of Lena.

Image	Our algorithm	Wang et al.'s algorithm [28]	Farhan and Sanjeev's algorithm [29]	Hong et al.'s algorithm [30]	
Lena	Horizontal	-0.0016	-0.0031	-0.0146	0.0020
	Vertical	0.0043	0.0084	0.0098	0.0042
	Diagonal	-0.0026	-0.0007	0.0056	0.0013

TABLE 2: Comparison of information entropy of ciphertext images.

Image	Our algorithm	Wang and Guan's algorithm [25]	Niyat et al.'s algorithm [26]	Wu et al.'s algorithm [27]
Lena	7.9988	7.9976	7.9974	7.9976
Peppers	7.9994	7.9980	7.9972	7.9974

TABLE 3: Comparison of NPCR (%) of ciphertext images.

Image	Our algorithm	Wang et al.'s algorithm [28]	Farhan and Sanjeev's algorithm [29]	Hong et al.'s algorithm [30]
Lena	99.6231	99.6016	99.6356	99.6037
Peppers	99.6307	99.6091	99.5891	99.6124

TABLE 4: Comparison of UACI (%) of ciphertext images.

Image	Our algorithm	Wang et al.'s algorithm [28]	Farhan and Sanjeev's algorithm [29]	Hong et al.'s algorithm [30]
Lena	33.4463	33.4735	33.4147	33.4381
Peppers	33.4768	33.6234	33.3568	33.7216

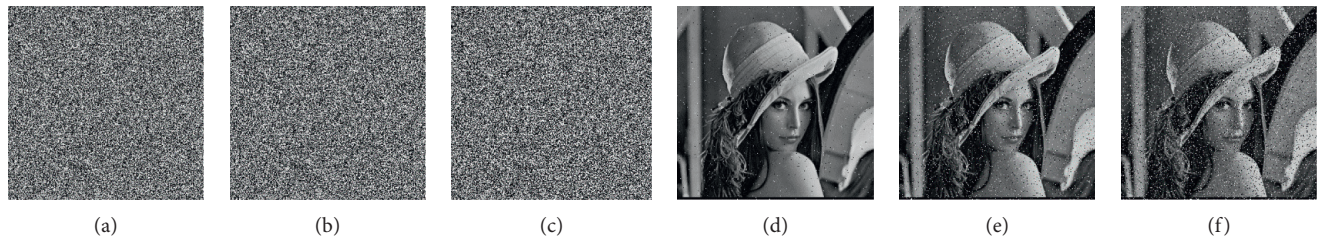


FIGURE 4: Noise test results. (a) Encrypted image of Lena with noise on the level of 0.01. (b) Encrypted image of Lena with noise on the level of 0.05. (c) Encrypted image of Lena with noise on the level of 0.1. (d) Decrypted image of Lena with noise on the level of 0.01. (e) Decrypted image of Lena with noise on the level of 0.05. (f) Decrypted image of Lena with noise on the level of 0.1.

Tables 3 and 4 compare the NPCR and UACI of the ciphertexts obtained by our algorithm with those of the ciphertexts obtained by three other algorithms. Compared with those of other algorithms, the NPCR and UACI of our algorithm were mostly approximately 99.6093% and 33.4635%, respectively.

**5.5. Robustness Analysis.** Robustness is an important indicator of the anti-disturbance ability of a cryptosystem. Robustness analysis means that the decryption algorithm can still decrypt the content of the image even when the image is disturbed by other information and means, and the corresponding information can be obtained through the decrypted image, so as to prove the robustness of the algorithm. The robustness of our algorithm was tested by noise attack and denial-of-service (DoS) attack. Noise interference is an important issue in actual communication. Common noises include Gaussian noise, salt-and-pepper

noise, etc. The salt-and-pepper noise stands out for its significant impact on ciphertext images. Therefore, this paper mainly tests the influence of the addition of salt-and-pepper noise to the plaintext image over our algorithm performance. Without changing the keys, different levels of salt-and-pepper noises were added to the plaintext image of Lena, and the noisy image was encrypted and decrypted by our algorithm. Figure 4 shows the encrypted and decrypted images of the plaintext image of Lena with salt-and-pepper noise on the level of 0.01, 0.05, and 0.1, respectively. Even when the noise level was 0.1, the plaintext image could be distinguished in the image decrypted by our algorithm, evidencing the strong resistance of our algorithm to noise attack.

The robustness of our algorithm was also tested against the DoS attack. Firstly, different portions of the information on the ciphertext image of Lena were blocked and then decrypted by our algorithm. Obviously, the quality of the decrypted image is negatively correlated with the amount of

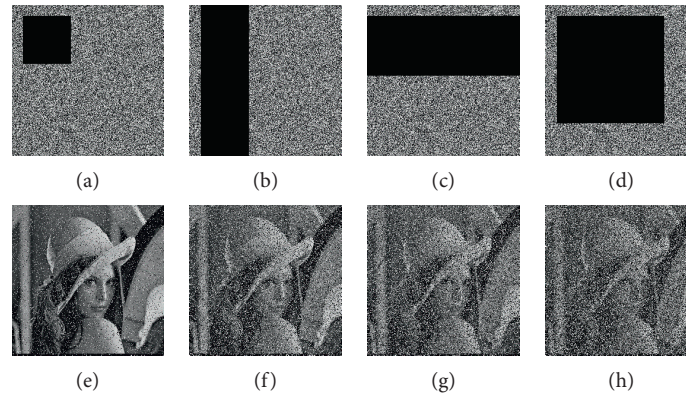


FIGURE 5: DoS attack test results. (a) Plaintext image of Lena with 6.25% of information being blocked. (b) Plaintext image of Lena with 25.44% of information being blocked. (c) Plaintext image of Lena with 40.20% of information being blocked. (d) Plaintext image of Lena with 87.89% of information being blocked. (e) Decrypted image of Lena with 6.25% of information being blocked. (f) Decrypted image of Lena with 25.44% of information being blocked. (g) Decrypted image of Lena with 40.20% of information being blocked. (h) Decrypted image of Lena with 87.89% of information being blocked.

information being blocked. Figure 5 presents the decrypted images of the ciphertext image of Lena with 6.25%, 25.44%, 40.20%, and 87.89% of information being blocked. It can be inferred that the main information of the plaintext image of Lena could still be recognized from the decrypted images. Therefore, our algorithm can effectively withstand the DoS attack and boasts strong robustness.

## 6. Conclusions

This paper mainly designs a chaotic image encryption algorithm based on Hopfield neural network and bidirectional flipping, a scrambling strategy. Firstly, the plaintext image was segmented into multiple blocks, and each block was scrambled. Then, SHA-512 hash algorithm was combined with the plaintext image to generate a hash array. On this basis, the initial values and control parameters were determined for the initialization of the chaotic mapping, and the control parameters were configured for scrambling and diffusion. Compared with other image encryption algorithms, our algorithm innovatively applies Hopfield neural network to encrypt images and adopts brand-new scrambling and diffusion models. Through simulation and theoretical analysis, it was confirmed that our algorithm is robust and effective in resisting statistical attack, differential attack, noise attack, and DoS attack.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This study was supported by Equipment Pre-Research Fund, General Armaments Department, People's Liberation Army

(grant no. 61421070101162107002), and General Project of Natural Science Foundation, Liaoning Province, China (grant no. 20170540426).

## References

- [1] W. Xue and Q. Lyu, "Image encryption algorithm based on Gray code and chaotic system," *Computer Systems & Applications*, vol. 27, no. 7, pp. 177–181, 2018.
- [2] Z. X. Jia and Y. P. Liu, "New image encryption algorithm based on adaptive and global scrambling," *Natural Science Edition*, vol. 2019, no. 6, pp. 61–72, 2019.
- [3] Y. Wang and L. Tu, "New image encryption algorithm based on improved Lorenz chaotic system," *Journal of Central South University: Natural Science Edition*, vol. 48, no. 10, pp. 2678–2685, 2017.
- [4] G. G. Bulut, M. C. Çatalbaş, and H. Güler, "Chaotic systems based real-time implementation of visual cryptography using LabVIEW," *Traitement du Signal*, vol. 37, no. 4, pp. 639–645, 2020.
- [5] C. J. Hu, C. Ruan, and Z. X. Niu, "Image encryption algorithm based on improved Logistic mapping," *Computer Systems & Applications*, vol. 28, no. 6, pp. 125–129, 2019.
- [6] M. Mollaefar, A. Sharif, and M. Nazari, "A novel encryption scheme for colored image based on high level chaotic maps," *Multimedia Tools & Applications*, vol. 76, pp. 1–23, 2017.
- [7] A. Mokhnache and L. Ziet, "Cryptanalysis of a permutation based image encryption technique using chaotic map," *Traitement du Signal*, vol. 37, no. 1, pp. 95–100, 2020.
- [8] J. Liu, Y. Ma, S. Li, J. Lian, and X. Zhang, "A new simple chaotic system and its application in medical image encryption," *Multimedia Tools and Applications*, vol. 77, no. 17, Article ID 22787, 2018.
- [9] D. Herbadji, N. Derouiche, A. Belmeguenai, A. Herbadji, and S. Boumerdassi, "A tweakable image encryption algorithm using an improved logistic chaotic map," *Traitement du Signal*, vol. 36, no. 5, pp. 407–417, 2019.
- [10] X. Chai, Y. Chen, and L. Broyde, "A novel chaos-based image encryption algorithm using DNA sequence operations," *Optics and Lasers in Engineering*, vol. 88, pp. 197–213, 2017.

- [11] X.-Y. Wang, Y.-Q. Zhang, and X.-M. Bao, "A novel chaotic image encryption scheme using DNA sequence operations," *Optics and Lasers in Engineering*, vol. 73, pp. 53–61, 2015.
- [12] A. Girdhar and V. Kumar, "A RGB image encryption technique using Lorenz and Rossler chaotic system on DNA sequences," *Multimedia Tools and Applications*, vol. 77, no. 20, Article ID 27017, 2018.
- [13] Z. Xia, X. Wang, W. Zhou, R. Li, C. Wang, and C. Zhang, "Color medical image lossless watermarking using chaotic system and accurate quaternion polar harmonic transforms," *Signal Processing*, vol. 157, pp. 108–118, 2019.
- [14] R. Zahmoul, R. Ejbali, and M. Zaied, "Image encryption based on new Beta chaotic maps," *Optics and Lasers in Engineering*, vol. 96, pp. 39–49, 2017.
- [15] C. Pak and L. Huang, "A new color image encryption using combination of the 1d chaotic map," *Signal Processing*, vol. 138, pp. 129–137, 2017.
- [16] S. Rajagopalan, S. Rethinam, and S. Arumugham, "Networked hardware assisted key image and chaotic attractors for secure RGB image communication," *Multimedia Tools and Applications*, vol. 77, pp. 1–34, 2018.
- [17] P. Praveenkumar, R. Amirharajan, and K. Thenmozhi, "Fusion of confusion and diffusion: a novel image encryption approach," *Telecommunication Systems*, vol. 65, no. 1, pp. 1–14, 2017.
- [18] R. Enayatifar, A. H. Abdullah, I. F. Isnin, A. Altameem, and M. Lee, "Image encryption using a synchronous permutation-diffusion technique," *Optics and Lasers in Engineering*, vol. 90, pp. 146–154, 2017.
- [19] X. Lu, X. Gou, and Z. Li, "A novel chaotic image encryption algorithm using block scrambling and dynamic index based diffusion," *Optics and Lasers in Engineering*, vol. 91, pp. 41–52, 2017.
- [20] X. P. Yan, X. Y. Wang, and Y. J. Xian, "Chaotic image encryption algorithm based on arithmetic sequence scrambling model and DNA encoding operation," *Multimedia Tools and Applications*, vol. 1, pp. 1–35, 2021.
- [21] A. Bakhshandeh and Z. Eslami, "An authenticated image encryption scheme based on chaotic maps and memory cellular automata," *Optics and Lasers in Engineering*, vol. 51, no. 6, pp. 665–673, 2013.
- [22] L. Liu, Q. Zhang, X. Wei, and C. Zhou, "Image encryption algorithm based on chaotic modulation of arnold dual scrambling and DNA computing," *Advanced Science Letters*, vol. 4, no. 11, pp. 3537–3542, 2011.
- [23] Q. Zhang, L. Guo, and X. P. Wei, "Image encryption using DNA addition combining with chaotic maps," *Mathematical & Computer Modelling an International Journal*, vol. 52, no. 11–12, pp. 2028–2035, 2010.
- [24] H. Liu, X. Wang, and A. Kadir, "Image encryption using DNA complementary rule and chaotic maps," *Applied Soft Computing*, vol. 12, no. 5, pp. 1457–1466, 2012.
- [25] X. Wang and N. Guan, "A novel chaotic image encryption algorithm based on extended Zigzag confusion and RNA operation," *Optics & Laser Technology*, vol. 131, Article ID 106366, 2020.
- [26] A. Yaghouti Niyat, M. H. Moattar, and M. Niazi Torshiz, "Color image encryption based on hybrid hyper-chaotic system and cellular automata," *Optics and Lasers in Engineering*, vol. 90, pp. 225–237, 2017.
- [27] J. Wu, X. Liao, and B. Yang, "Image encryption using 2D Hénon-Sine map and DNA approach," *Signal Processing*, vol. 153, pp. 11–23, 2018.
- [28] X. Y. Wang, J. J. Zhang, and F. C. Zhang, "New chaotical image encryption algorithm based on Fisher-Yates scrambling and DNA coding," *Chinese Physics B*, vol. 28, no. 4, pp. 125–134, 2019.
- [29] M. Farhan and K. Sanjeev, "A novel fractional order chaos-based image encryption using Fisher yates algorithm and 3-D cat map," *Multimedia Tools and Applications*, vol. 78, no. 11, Article ID 14867, 2019.
- [30] M. Y. Hong, L. Ye, and H. G. Li, "A new image cryptosystem based on 2D hyper-chaotic system," *Multimedia Tools and Applications*, vol. 15, no. 76, pp. 8087–8108, 2017.



## Retraction

# Retracted: Visualization and Analysis of Mapping Knowledge Domain of Heterogeneous Traffic Flow

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] Y. He, Q. Feng, L. Yan, and X. Lu, "Visualization and Analysis of Mapping Knowledge Domain of Heterogeneous Traffic Flow," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 7754961, 15 pages, 2022.

## Research Article

# Visualization and Analysis of Mapping Knowledge Domain of Heterogeneous Traffic Flow

Yi He <sup>1</sup>, Qi Feng <sup>1</sup>, Lixin Yan <sup>2</sup> and Xiao-Yun Lu<sup>3</sup>

<sup>1</sup>Intelligent Transport Systems Research Center, Wuhan University of Technology, Wuhan 430063, China

<sup>2</sup>School of Transportation and Logistics, East China Jiaotong University, Nanchang 430063, China

<sup>3</sup>California PATH, University of California, Berkeley, CA 94804-4684, USA

Correspondence should be addressed to Lixin Yan; yanlixinits@163.com

Received 13 December 2021; Accepted 30 December 2021; Published 4 February 2022

Academic Editor: Daqing Gong

Copyright © 2022 Yi He et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Mapping knowledge domain (MKD) is an important application in bibliometrics, which is a method of visually presenting and explaining newly developed interdisciplinary scientific fields using data mining, information analysis, scientific measurement, and graphic rendering. This study combines applied mathematics, visual analysis technology, information science, and scientometrics to systematically analyze the development status, research distribution, and future trend of the heterogeneous traffic flow by using the MKD software tools VOSviewer and CiteSpace. Based on the MKD and Bibliometrics approaches, 4709 articles have been studied, which were published by Science Citation Index Expanded (SCIE) and Social Sciences Citation Index (SSCI) from 2004 to 2021 in the field of heterogeneous traffic flows. Firstly, this paper presents the annual numbers of articles, origin countries, main research organizations, and groups as well as the source journals on heterogeneous traffic flow studies. Then, cocitation analysis is used to divide heterogeneous traffic flow into three main research directions, which include “heterogeneous traffic flow model,” “traffic flow capacity analysis,” and “traffic flow stability analysis.” The keyword cooccurrence analysis is applied to identify five dominant clusters: “modeling and optimization methods,” “traffic flow characteristics analysis,” “driving behavior analysis,” “simulation experiment,” and “policies and barriers.” Finally, burst keywords were studied according to the publication date to present more clearly the change of research focus and direction over time.

## 1. Introduce

As the production and ownership of automobiles increase year by year, vehicles not only provide convenience and comfort to people but also bring many negative impacts, such as traffic safety, congestion, and air pollution [1, 2]. Vehicle crashes result in about 63,194 fatalities in China, and direct property losses reached 11.18672 million yuan in 2018 [3]. In terms of energy consumption, China’s road transportation industry consumes 60.1% of the nation’s oil and produces 14% of greenhouse gases [4], while America’s transportation industry consumes 70% of oil and 33.6% of greenhouse gases [5]. In the face of increasingly severe security, energy, and environmental problems, autonomous driving technology, as an important component of the intelligent transportation system, has incomparable

advantages in improving travel efficiency, alleviating traffic congestion, and reducing gasoline consumption and other aspects and is an important way to achieve green and efficient development of the automobile industry.

Like many new technologies, although some autonomous driving technologies have been implemented, most of them are still in the experimental and conceptual stage, and the industrialization of autonomous driving vehicles still needs a long process. Boston Consulting Group predicts that “it will take 15–20 years for the global market share of autonomous vehicles to reach 25%,” and as autonomous vehicles are expected to hit the market in 2021, this means that autonomous vehicles will account for 25% of the global market in 2035–2040 [6]. Therefore, in the process of the rapid development of the autonomous vehicle industry, there will be a transitional stage of the heterogeneous



transportation system composed of artificial and autonomous vehicles [7].

Heterogeneous traffic flow is an extremely complex problem in the field of traffic flow, which involves the combination of traditional traffic flow research and autonomous driving technology. Although some article reviews have studied the related content of heterogeneous traffic flow, previous article reviews have adopted qualitative rather than quantitative methods to conduct research, which largely relies on the author's experience and subjective judgment in this field, requiring the author to accumulate, summarize, and refine in this field for a long time. In addition, the previous review article lacks visual analysis of all heterogeneous traffic flow studies, so it is difficult to reveal the panorama of heterogeneous traffic flow research, and it is difficult to describe the cocitation relationship and keyword evolution law between articles [8].

In this paper, the bibliometrics method was used to review the article. The WOS platform was used to search 4709 related articles published in the field of heterogeneous traffic flow in the recent 17 years (2004–2021). The following tasks were completed: which countries/regions are the most important research results; what is the collaboration between the main authors in the research area; what are the core publications in the research area; what are the most influential research results in the field of heterogeneous traffic flow research; which are the most active research field of heterogeneous traffic flow and its present situation and deficiency? Research and review related aspects from a new perspective clarify the development context of the heterogeneous traffic flow field, identify key topics and research progress in domestic and foreign research, find the shortcomings of current research, and provide reference and comparison for researchers.

The rest of the paper is organized as follows: Section 2 introduces the data sources of the sample literature as well as the analytical tools and methods; Section 3 carries out quantitative statistical analysis from aspects of publication year, country, institution, journal, and keywords through sorting out 4709 high-level sample literature related to heterogeneous traffic flow; Section 4 draws conclusions and provides a brief discussion of the future research plan.

## 2. Data Sources and Analysis Methods

**2.1. Data Source.** This article aims to conduct a quantitative and visualized analysis of the representative works related to heterogeneous traffic flow through MKD methods. Web of Science (WOS) is an online subscription-based science citation index service that provides access to multiple databases containing more than 148 million records (journals, books, and conference proceedings) dating back to the beginning of the 20th century [9]. Therefore, the SCIE and SSCI citation index database in the WOS Core Collection were retrieved as the source in this study. To obtain the maximum number of relevant documents related to simulated driving, we collected and stored data for the defined search terms using the “Topic” search in Web of Science Core Collection. The keywords used for the initial data

collection included “heterogeneous traffic flow\*” OR “Mixed traffic flow\*” OR (“Penetration rate\*”) AND (“connected and autonomous vehicle\*” OR “intelligent connected vehicle\*” OR “connected vehicle\*” OR “autonomous vehicle\*” OR “adaptive cruise control” OR “cooperative adaptive cruise control”), where “\*” represents fuzzy search.

The excluded categories were veterinary medicine, clinical medicine, food science, engineering, etc., with a time span of 1985 to 2021 and retrieval date of January 2, 2021, 754 English journal articles, conference papers, research reviews, and conference abstractions were obtained. To ensure the completeness and accuracy of the article collection, this paper added the cited articles of 754 articles obtained from article retrieval and removed the remaining 3955 articles after self-cited articles. Finally, a total of 4709 articles were obtained for article analysis, involving 7046 authors and 13,224 keywords. To cover as much additional information as possible, the “full records and references” of all articles are exported in “TXT” format for bibliometrics analysis.

**2.2. Analysis Method.** Bibliometrics is an interdisciplinary science that uses mathematical and statistical methods to quantitatively analyze relevant articles, mainly studying the quantitative relationship of the article (various publications, in particular, journal papers and citations are in the majority) [10], journal information sources (the situation of journals in which article is published) [11], author relationship (the cooperative relationship between the author, the organization, or the group) [12], and keywords (the core words that repeatedly appear in the article) [13].

In bibliometrics analysis, the mapping knowledge domain (MKD) is a method to visually present and explain the newly developed interdisciplinary scientific field using data mining, information analysis, scientific measurement, and graph rendering, aiming at simplifying the information acquisition process in the research field and clarifying the knowledge structure [14]. The MKD has dual properties and characteristics of “graph” and “spectrum,” which means that it is not only a visual knowledge map but also a serialized knowledge map. The MKD can show the development process and structural relations of scientific knowledge as well as the implied complex relations, including network, structure, interaction, intersection, evolution, or derivation between knowledge units or knowledge clusters [15]; understanding these complex knowledge relationships can produce new knowledge.

The mapping (or creation) of MKD includes article cocitation analysis, keyword cooccurrence analysis, and burst detection analysis, as described as follows:

- (a) Document cocitation analysis: a document cocitation is a new form of article coupling, which refers to the frequency of two articles being cited simultaneously [16]. Document cocitation analysis refers to the statistics of the number of two documents simultaneously cited by one or more documents, to conduct network analysis and cluster analysis on the

cited documents, to analyze the knowledge base of the specific subject represented by these documents [17]. The knowledge base of this field is constructed by quoting articles, and the research frontier of a certain field is formed by analyzing the knowledge base.

- (b) Keywords cooccurrence analysis: official academic publications are generally accompanied by keywords, which not only reflect the core content of the research results and condense the author's academic views but also provide an important way of retrieval. Keywords cooccurrence analysis is to count the number of occurrences of a group of keywords in the same article in pairs [17]. Based on this, network analysis and cluster analysis are conducted for these words, to reflect the affinity relationship between these words and the structural changes of the subject they represent [18]. Therefore, the more frequently a keyword appears, the more relevant research results, and the stronger the concentration of research content [19].
- (c) Burst detection analysis: the burst detection analysis considers the change of keyword frequency and determines the keywords with the characteristics of burst growth in a certain period in a certain research field, which can be used to study the development trend of a certain topic. It is different from the burst detection analysis based on a threshold value. Although the frequency of each keyword may be relatively low, the burst keyword can be found according to the change of keyword frequency, so the latest research trends can be predicted by such keywords.

**2.3. Analysis Tools.** In this study, the method of MKD in bibliometrics was selected to analyze the article according to its attributes and characteristics, and the knowledge was visualized. CiteSpace and VOSviewer were used in the study. CiteSpace was developed by Professor Chen Chaomei, chair professor of Changjiang Scholars in Dalian University of Technology [20]. The software focuses on the potential knowledge contained in scientific analysis. At first, it specifically analyzes the cocitation of articles and excavates the knowledge clustering and distribution in citation space. In subsequent updates, cooccurrence analysis between other knowledge units will also be provided, such as author, institution, and country or region collaboration. VOSviewer software was developed by Van Eck and Waltman of Leiden University in the Netherlands. Visualization of Similarities (VOS) technology is used to perform a visualization analysis of article knowledge units. Cooccurrence matrix is used to generate a knowledge map. Its core principles include the construction of similarity matrix and the VOS layout method [21].

**2.4. Construction of MKD.** By normalizing the cooccurrence matrix, the similarity matrix was obtained by correcting the differences in the total occurrence times or common

frequency of elements in the cooccurrence matrix (i.e., author, institution, country/region, keywords, etc.). VOSviewer uses association strength, also known as proximity index, to measure the similarity between pairs of elements in the cooccurrence matrix. The similarity  $S_{ij}$  between two items  $i$  and  $j$  is calculated as

$$S_{ij} = \frac{c_{ij}}{w_i w_j}, \quad (1)$$

where  $S_{ij}$  denotes the similarity between the elements  $i$  and  $j$ ,  $c_{ij}$  denotes the number of times the element  $i$  and the element  $j$  cooccur, and  $w_i$  and  $w_j$  denote the total number of occurrences of the element  $i$  and the element  $j$ , respectively. It can be seen that the normalized value is between 0 and 1, where the larger the value, the higher the similarity between elements, and the smaller the value, the smaller the similarity between elements.

**2.5. Layout Method.** The layout method adopted in this paper is to reflect the similarity between any pair of elements  $i$  and  $j$  by the spatial distance  $s_{ij}$  to determine the position of elements in two-dimensional space. The closer the distance between elements, the higher the similarity [22]. To make the clustering effect more obvious, the main idea of VOS is to minimize the sum of the weighted Euclidean distances of all objects in each cluster; the formula is as follows [21]:

$$D(x_1, \dots, x_n) = \sum_{i < j} s_{ij} x_i - x_j^2, \quad (2)$$

where  $n$  is the number of elements to be laid out,  $\|\bullet\|$  denotes the Euclidean norm, and vector  $x_i = (x_{i1}, x_{i2})$  denotes the location of item  $i$  in a two-dimensional map. To avoid all elements being in the same position, the objective function is minimized under the following constraints:

$$\frac{2}{n(n-1)} \sum_{i < j} x_i - x_j = 1. \quad (3)$$

### 3. Article Statistical Analysis

**3.1. Annual Distribution Statistics of the Article.** The annual distribution of articles refers to the difference in the distribution of the number of articles in the same time zone, which can reflect the development process of people, institutions, or subject research and divide the active trend of subject research. The change in the number of academic papers on a subject is an important indicator of the development trend of the research field, and it is also a reflection of the change in the scope of subject knowledge. By drawing a graph of the number of documents and conducting multivariate statistical analysis, one can understand the research level and future development trend of a certain field.

To better explore the relevant characteristics and conditions in the field of heterogeneous traffic flow in a period, this paper considers the number of documents in the field of heterogeneous traffic flow, which are divided into three

stages based on the growth of the article during the period, namely, “initial stage (2004–2009),” “preliminary development stage (2010–2015),” and “rapid development stage (2016–2020).” On the one hand, segmenting the field according to the same time can better compare and analyze the growth of articles and the characteristics of research trends in each period. On the other hand, because the research on heterogeneous traffic flow started late and the period is not long, it is relatively reasonable to set 5 years as an academic research cycle.

According to the search results, the earliest research article on heterogeneous traffic flow was published in 2004 and lasted until 2021. It already contains 4703 related documents. The annual distribution statistics are shown in Figure 1. It is worth noting that due to 2020 the data is still being updated, so the data in the statistics is small.

**Initial stage (2004–2009):** from the first article published in 2004 to 2008, there were few related research results in this field, and only 73 papers (2009) were published each year, which means that there is research on qualitative traffic flow has just started, and a complete document system has not yet been formed.

**Primary development stage (2010–2015):** as countries pay more and more attention to road safety, autonomous driving has set off a research boom and has begun to enter the primary development stage. The number of relevant documents at this stage has increased significantly, with an average increase of 33 documents per year. It can be considered that the field of heterogeneous traffic flow research was initially formed during this period.

**Rapid development stage (2016–2020):** after the accumulation of knowledge in the initial stage and the primary development stage, the number of relevant documents in this stage has grown rapidly. This shows that the research enthusiasm in the field of heterogeneous traffic flow is continuously increasing and has entered a stage of rapid development.

**3.2. Country and Regional Statistics.** For national and regional factors, this article presents four indicators to analyze the distribution area and situation of the documents, namely, the number of documents (articles), the percentage of documents in the total number of documents, the total number of citations of documents (times), and the average single document. The number of citations is shown in Table 1. The use of these indicators mainly considers two aspects: on the one hand, the number of documents in a country can reflect the research enthusiasm of researchers in each country in the field; on the other hand, the total number of citations and the average number of citations per article can be to a certain extent. This indicates the quality of the article of national researchers, reflecting the depth of research in this field by researchers from various countries.

According to the retrieved results, the article on heterogeneous traffic flow mainly comes from 88 countries (or regions). Table 1 lists the top 10 countries in the field of heterogeneous traffic flow research from 2004 to the present. Japanese science historian Yuasa (1962) defined a country

with more than 25% of the world’s major scientific achievements as the center of scientific activity. According to the data in Table 1, China ranks first with 2087 articles, accounting for 44.28% of the total number of articles. The average number of citations per article is 8.27, followed by the United States (1139) and India (342). The center of scientific activity in the field of heterogeneous traffic flow research is as follows. Four Asian countries (China, India, Japan, and South Korea) among the top ten countries in the number of articles indicate that Asian countries have performed well in research on heterogeneous traffic flow. European countries have excellent performance in the average number of citations per article, such as Germany (average number of citations per article 15.49), the United Kingdom (average number of citations per article 14.61), and the Netherlands (average number of citations per article 13.57), which shows that the article published in these countries has been highly recognized by researchers.

In VOSviewer, citation analysis is used to generate a scientific knowledge map of the country in the field of heterogeneous traffic flow. This article uses VOSviewer software to analyze the collected data and make a density view based on the distribution of national documents. The density view can visually display the situation of the country/region in the density dimension. Each label in the figure represents a country, and the label is the center point to form a region. The size and color of the region depend on the country/region represented by the label the number of published documents, the larger the number, the brighter the color, and the smaller the number, the darker the color. The location of the tag depends on the number of items near the tag and their importance. The density view can help understand the overall structure of the map and draw attention to the most important areas of the map.

This article sets the display conditions for countries/regions with five or more related article achievements, a total of 60 countries/regions, as shown in Figure 2. The density view can see the overall structure of the distribution of national or regional documents. In the figure, China and the United States are in the center of the figure, indicating that China and the United States play an important role in the field of heterogeneous traffic flow.

**3.3. Analysis of Main Research Institutions.** By analyzing the cooperation relationship between research institutions, information about the most outstanding organizations and groups in a certain field in the discipline can be determined. To analyze the cooperative relationship between the research institutions related to the article, this paper uses VOSviewer to construct a cooperation network between major research institutions to analyze the participating institutions of heterogeneous traffic flow research, as shown in Figure 3.

In the interinstitutional cooperation network diagram, the label is the name of the research institution, the size of the node indicates the number of standardized documents, and the connection indicates the cooperative relationship between the two connected research institutions. The closer the cooperation, the closer the node distance, the wider the

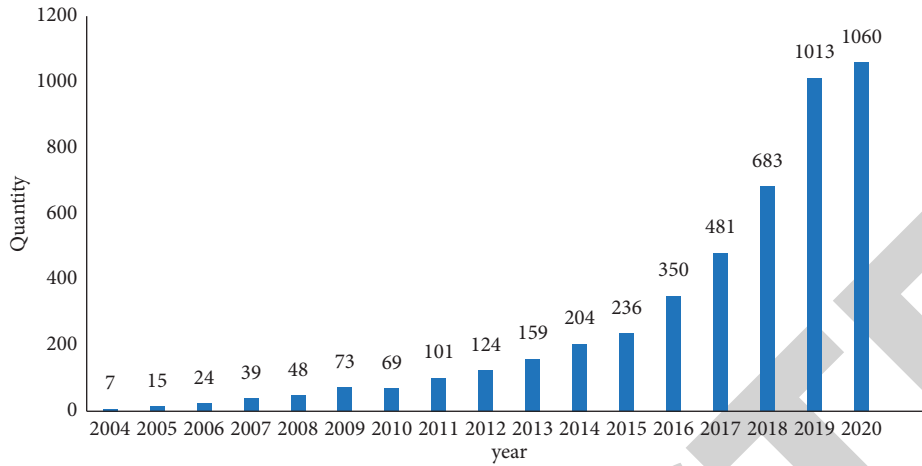


FIGURE 1: Quantitative distribution of published articles in heterogeneous traffic flow studies, 2004–2020.

TABLE 1: The top ten countries in the number of articles in the heterogeneous traffic flow research field.

Rank	Country/territories	Quantity	Percentage	Total cited times	The average number of citations per article
1	China	2087	44.28	17251	8.27
2	USA	1139	24.17	11518	10.11
3	India	342	7.26	2894	8.46
4	Germany	206	4.37	3190	15.49
5	England	187	3.97	2732	14.61
6	Australia	167	3.54	1906	11.41
7	Netherlands	152	3.23	2063	13.57
8	Canada	139	2.95	1283	9.23
9	Japan	131	2.78	1122	8.56
10	South Korea	121	2.57	720	5.95

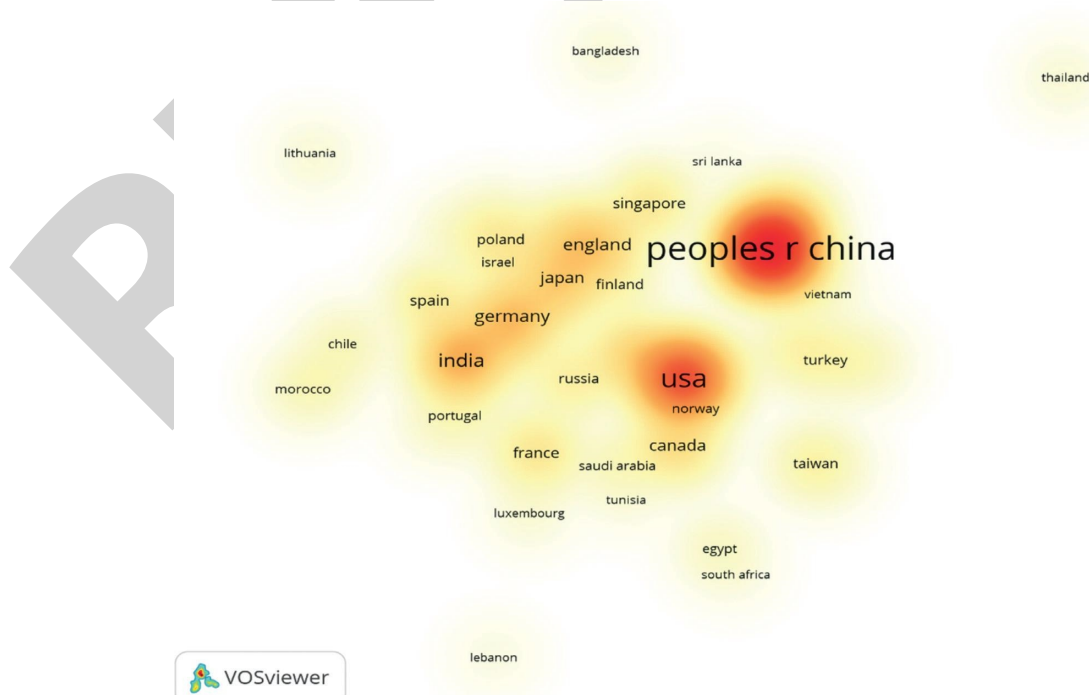


FIGURE 2: Density distribution of heterogeneous traffic flow in major countries.

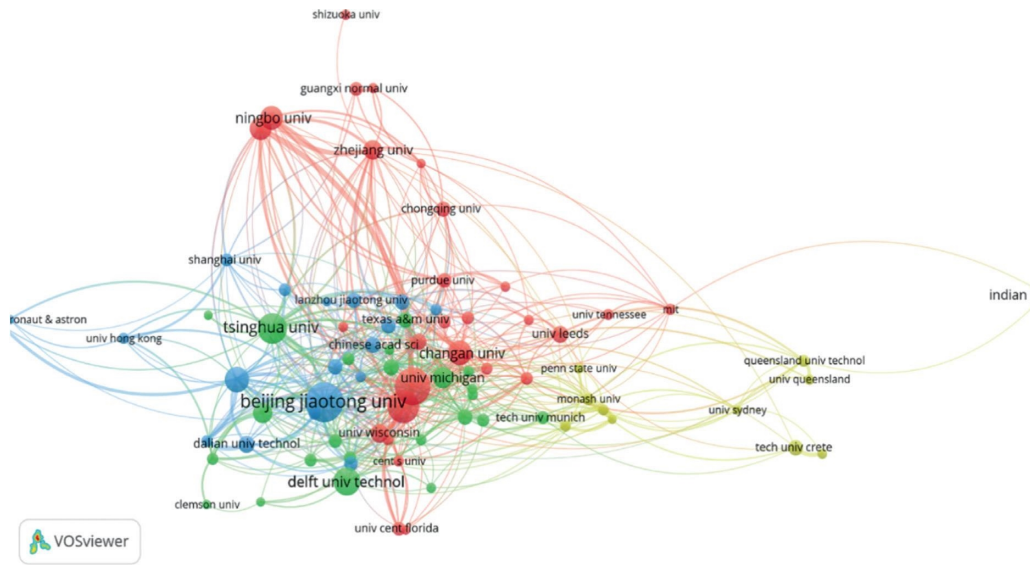


FIGURE 3: Collaborative networks among major research institutions.

scope of cooperation, the more connections. Figure 3 shows that Beijing Jiao tong University is at the center of the gathering point, indicating that Beijing Jiao tong University and other research institutions have established close cooperation in the field of heterogeneous traffic flow.

The number of published documents of research institutions is also one of the important indicators to measure the contributions made by research institutions in the field. The article in this field comes from a total of 2512 organizations. Table 2 lists the top 10 organizations with the largest number of published articles. The research strength of a university is often an indicator of the national scientific research and innovation ability. Table 2 shows that a total of 9 universities contributed 1173 papers, accounting for 24.89% of the total number of heterogeneous traffic flow publications. Among them, Chinese universities and research institutes are heterogeneous traffic. The main contributors to stream research are as follows: these organizations have published 1059 papers, accounting for 22.47% of the total number of papers. Among Chinese universities, Beijing Jiaotong University ranked first in the number of publications of 213 articles, with an average of 161.85 citations per article, indicating that Beijing Jiaotong University is in a leading position in the field of heterogeneous traffic flow research. Southeast University (187 articles) and Tongji University (154 articles) followed closely. Among the 2512 organizations, only 87 organizations published 20 papers or more, accounting for 3.46% of the total. This result shows that the published documents are concentrated in a few organizations.

**3.4. Coauthor Analysis of the Main Research Group.** Counting the number of articles published by authors on the topic and comparing and analyzing them can understand the main authors and core author groups in the field and can further establish their important role and influence in related

fields by understanding the author's research progress and by paying attention to them. The research direction and focus, to understand its leading and leading role in the development of the whole discipline, is of great significance for further understanding the current situation and future development of special research. Because the research in the field of heterogeneous traffic flow is highly interdisciplinary, researchers come from different fields such as traffic engineering, traffic planning, mathematics, psychology, computer science, medicine, and statistics, and cooperation can achieve complementary advantages. The creation and analysis of the knowledge graph of the coauthor network of the article can provide valuable information for research institutions to develop cooperation groups and seek cooperation opportunities for individual researchers and for publishers to form editorial teams (publishing books or special issues in journals). In VOSviewer, use coauthor analysis to generate a knowledge domain map of the coauthors of the main research group.

In the coauthor relationship diagram of major research groups, the label is the abbreviation of the author's name, the node size represents the number of articles published by the author after standardization, and the line represents the cooperative relationship between the two authors connected. The closer the cooperation, the closer the node distance, the wider the scope of cooperation, and the more links. According to VOSviewer's calculations, there are 7046 authors involved in the heterogeneous traffic flow field. To clearly show the cooperative relationship between authors, the minimum number of published papers of authors is set as 5, the minimum number of citations is set as 100, and a total of 224 authors are displayed at last. Figure 4 reflects the visualization of authors' cooperative relationship network in the heterogeneous traffic flow field, which shows a relatively large number of authors' achievements with the close cooperative relationships. The results show that the

TABLE 2: Organization of the top 10 research achievements in heterogeneous traffic flow from 2004 to 2021.

Rank	Organization	Link strength	Quantity	Percentage	Cited times	The average number of citations per article
1	Beijing Jiaotong University	125	213	4.52	2276	161.85
2	Southeast University	129	187	3.97	1116	143.58
3	Tongji University	106	154	3.27	852	142.86
4	Tsinghua University	86	133	2.82	1200	147.99
5	Delft University technology	24	120	2.55	1874	192.28
6	Beihang University	80	100	2.12	1803	169.31
7	Changan University	52	93	1.97	456	86.69
8	Ningbo University	103	92	1.95	1013	128.08
9	Jiangsu Provincial Collaborative Innovation Center modern urban	109	87	1.85	892	110.16
10	Indian Institute of Technology	4	81	1.72	1106	66.43

strengthening of cooperative relationships between authors can improve research efficiency and promote the development of heterogeneous traffic flow fields.

**3.5. Source Journal Analysis.** The source journal is an important information carrier for the dissemination and inheritance of scientific achievements. The analysis of journals in the academic field will determine the distribution of core journals in the field. According to the retrieved results, 4709 articles were published in 1095 journals, covering engineering, psychology, transportation, computer science, and other research fields.

Table 3 lists the top ten journals with the number of publications of heterogeneous traffic flow. "Transportation research part c-emerging technologies" publishes the largest number of documents in the field of heterogeneous traffic flow, up to 283 articles, with a total of 6129 citations. The article has been cited 21.66 times, with an impact factor of 6.077. The journal mainly deals with the development, application, and impact of transportation systems and emerging technologies, especially the impact of emerging technologies on the performance of transportation systems in terms of monitoring, efficiency, safety, and reliability. "Physical a-statistical mechanics and its applications" has the second largest number of publications, with a total of 265 articles and a total of 4485 citations, with an average of 16.92 citations per article. The journal's impact factor is 2.924. "Transportation Research Record" ranked third in the number of articles published, with a total of 235 articles, with a total of 1754 citations, with an average of 7.46 citations per article.

**3.6. Cocitation Analysis of Documents.** The document cocitation analysis was proposed by Henry Small in 1973. Its main principle is to measure the similarity between the documents by the number of times two documents are cited together; that is, if two documents are cited by the third document at the same time, the first two documents constitute a "cocitation" relationship, and their cocitation relationship increases with the increase in the number of citations. Document cocitation analysis is often used to explore the internal connections of scientific documents and

describe the dynamic structure of scientific development. Document cocitation network shows the spatial location of the most cited results in the research field in the form of graphs.

This section screens all the documents with the minimum number of citations of 40. A total of 149 documents meeting the threshold condition are extracted from 4709 documents. The visualization results of the document cocitation view are shown in Figure 5. The document cocited knowledge graph shows the scientific knowledge base and research frontiers. The size of the node in the graph indicates the total frequency of a document being cited. The larger the node, the higher the cited frequency and the greater the influence of the document; the color of the node indicates the document. Belonging to the cluster, the documents of the same cluster have greater similarity in the research topic. As shown in Figure 5, the entire knowledge network map is divided into three color-coded clusters using the default clustering approach:

- (1) Cluster 1 (Red): traffic capacity: in this cluster, the classic article with the largest density is "Enhanced Intelligent Driver Model to Access the Impact of Driving Strategies on Traffic Capacity" by Kesting et al. published in philosophical transactions of the royal society A. The number of cocitations is 283 and the total link strength is 24, indicating that this paper plays an important role in the structure of the cocitations network. In this article, Kesting investigates the influence of variable percentages of ACC vehicles on traffic flow characteristics through simulation and proposes a new car-following model that is based on the intelligent driver model (IDM) and inherits its intuitive behavioral parameters: desired velocity, acceleration, comfortable deceleration, and desired minimum time headway. At the same time, he eliminates some unrealistic behavior of the IDM. Simulation results show that sensitivities of the order of 0.3, i.e., 1 percent more ACC vehicles will lead to an increase in the capacities by about 0.3 percent with a suitable strategy. This sensitivity multiplies when considering travel times at actual breakdowns [23].



Another important article is “Impacts of Cooperative Adaptive Cruise Control on Freeway Traffic Flow” by Shladover et al. published in *Transportation Research Record: Journal of the Transportation Research Board*. The number of cocitations is 247 and the total link strength is 11. This study estimates the effect on highway capacity of varying market penetrations of vehicles with adaptive cruise control (ACC) and cooperative adaptive cruise control (CACC). The results showed that the use of ACC was unlikely to change lane capacity significantly [24].

- (2) Cluster 2 (Blue): traffic flow model: in this cluster, the classic article with the largest density is “Short-Term Traffic Forecasting: Where We Are and Where We’re Going” by Vlahogianni et al. published in the *Transportation Research Part C: Emerging Technologies*. The number of cocitations is 418 and the total link strength is 4. In this article, ten challenging but relatively under-studied directions of short-term traffic forecasting models are presented, and the existing challenges are summarized to make recommendations for future work [25].

In this field, we also recommend “A Markov Model for Headway/Spacing Distribution of Road Traffic” and “Three-Phase Traffic Theory and Two-Phase Models with a Fundamental Diagram in the Light of Empirical Stylized Facts” as two important articles according to the ranking.

“A Markov Model for Headway/Spacing Distribution of Road Traffic” by Chen and Li is published in *IEEE Transactions on Intelligent Transportation Systems*. The authors link two research directions of road traffic—the mesoscopic headway distribution model and the microscopic vehicle interaction model—together to account for the empirical headway/spacing distributions. In this model, empirical headway/spacing distributions are viewed as the outcomes of stochastic car-following behaviors and the reflections of the unconscious and inaccurate perceptions of space and/or time intervals that people may have [26].

“Three-Phase Traffic Theory and Two-Phase Models with a Fundamental Diagram in the Light of Empirical Stylized Facts” by Treiber et al. is published in *Transportation Research Part B*. This article compares Kerner’s three-phase traffic theory with the phase diagram approach for traffic models with a fundamental diagram and demonstrates that models created to reproduce three-phase traffic theory create similar spatiotemporal traffic states and associated phase diagrams, no matter whether the parameters imply a fundamental diagram in equilibrium or nonunique flow-density relationships [27].

- (3) Cluster 3 (Green): traffic flow stability analysis: traffic jams are usually related to the instability of traffic flow. Two kinds of traffic flow models can be used for stability analysis: the micro model and the macro

model. In this cluster, “Analytical Studies on the Instabilities of Heterogeneous Intelligent Traffic Flow” by Ngoduy uses microscopic models, particularly the effect of intelligent vehicles on heterogeneous (or multiclass) traffic flow instabilities. The analytical results show that time delay destabilizes traffic flow as found in the article and that the higher intelligent vehicle percentages, the more stable traffic flow with respect to a small perturbation for a given model parameter set [28].

In this field, we also recommend “Analysis on Traffic Stability and Capacity for Mixed Traffic Flow with Platoons of Intelligent Connected Vehicles” and “Longitudinal Emissions Evaluation of Mixed (Cooperative) Adaptive Cruise Control Traffic Flow and Its Relationship with Stability” as two important articles.

“Analysis on Traffic Stability and Capacity for Mixed Traffic Flow with Platoons of Intelligent Connected Vehicles” by Xin Chang and Haijian Li (2020) is published in *Physica A*. In this paper, analytical methods for the stability and the fundamental diagram models of mixed traffic flow are developed, and the results of the sensitivity analysis demonstrated that ICVs can improve the stability of the mixed traffic flow at a critical speed. However, if this critical speed is exceeded, an increase in the number of ICVs may degrade the stability of the mixed traffic flow [29].

“Longitudinal Emissions Evaluation of Mixed (Cooperative) Adaptive Cruise Control Traffic Flow and Its Relationship with Stability” by Yanyan Qin (2020) shows the impacts of the mixed CACC-MDV traffic on fuel consumption and emissions, by taking into consideration partial degenerations from stable CACC vehicles to unstable ACC vehicles. The results show that stability situations of the mixed traffic qualitatively influence the impact trend of CACC MPRs on fuel consumption and emissions [30].

**3.7. Keyword Cooccurrence Analysis.** Keyword cooccurrence analysis is a common research method in bibliometrics. By studying the cooccurrence relationship of cooccurring keywords in a large number of documents, it is used to analyze the link strength between cooccurring keywords. Its usage is to describe the internal relationship and structure of a certain academic field and to reveal the research front of the subject. The research front refers to the conceptual combination of temporary research topics and basic research issues, as well as theoretical trends and new topics that arise or emerge unexpectedly. In VOSviewer, cooccurrence analysis is used to generate a heterogeneous traffic flow research cooccurring keyword network, as shown in Figure 6.

In Figure 6, it can be seen that the frontier topics of heterogeneous traffic flow research form five groups, and the keywords in the same group show greater similarity in the research topics. According to the characteristics and current situation of heterogeneous traffic flow research, the following five categories are analyzed:

TABLE 3: The top ten journals of heterogeneous traffic flow publications.

Rank	Journal	Impact factor	Quantity	Cited times	The average number of citations per article
1	Transportation Research Part C: Emerging Technologies	6.077	283	6129	21.66
2	Physica A: Statistical Mechanics and Its Applications	2.924	265	4485	16.92
3	Transportation Research Record	1.56	235	1754	7.46
4	IEEE Transactions on Intelligent Transportation Systems	1.029	140	2031	14.51
5	IEEE Access	3.745	132	473	3.58
6	Journal of Advanced Transportation	1.67	129	598	4.64
7	Transportation Research Part B: Methodological	4.796	97	1785	18.40
8	Acta Physica Sinica	0.624	84	1008	12.00
9	Sustainability	2.576	75	140	1.87
10	Accident Analysis and Prevention	3.655	72	668	9.28

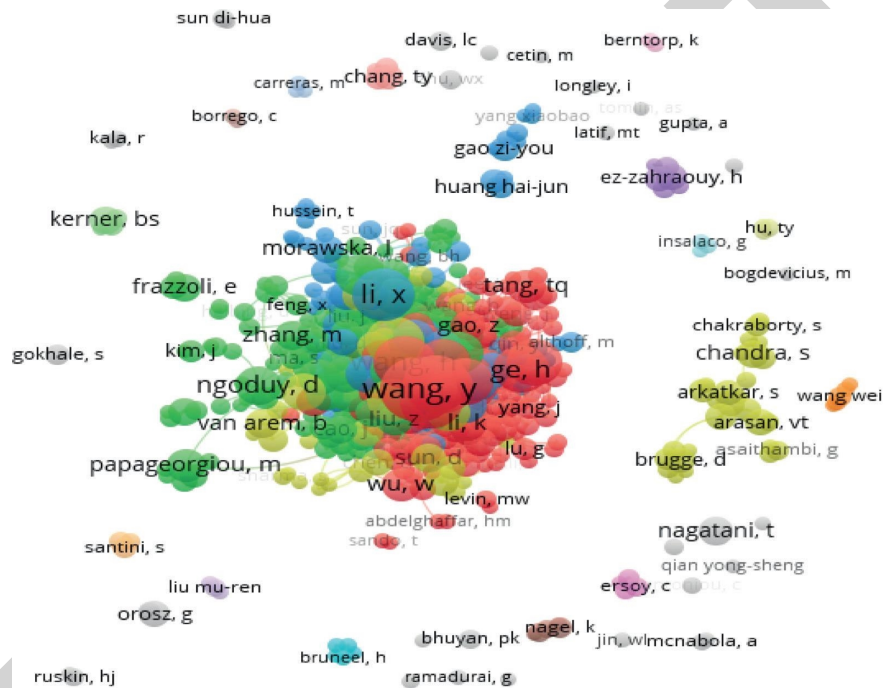


FIGURE 4: Coauthor relationships among major research groups.

3.7.1. *Group 1 (Red): Analysis of Modeling and Optimization Methods for Heterogeneous Traffic Flow.* This cluster contains a total of 55 keywords, mainly including model, systems, algorithm, optimization, and framework. The visualization results of the cluster show that keywords such as model, algorithm, and optimization are cooccurring high-frequency words.

From the perspective of research, the traffic flow model can be divided into a macro traffic flow model and a micro traffic flow model [31]. The microscopic traffic flow model processes traffic flow into dispersed particles to reflect the characteristics of traffic flow [32–35]. The macroscopic traffic flow model regards all vehicles on the whole road as a compressible continuous fluid, deduces the relationship among vehicle density  $\rho$ , the average velocity  $V$ , and flow  $Q$ , and thus describes the mathematical model of traffic flow operation law [36–39]. So far, the researchers have developed a variety of traffic models used to study the complex

phenomenon and obtained many important results [40, 41]; however, these traditional traffic flow models did not consider the information interaction between the vehicle and synergy, also lack the research of vehicle heterogeneous factors, and therefore, cannot be directly used to study the heterogeneous traffic flow.

To study the heterogeneity of traffic flow, Tang et al. [42] proposed a new vehicle following model based on the properties of heterogeneous traffic flow and the relationship between micro and macro variables, establishing a new dynamic model for heterogeneous traffic flow. Ngoduy et al. [43] established a new macro traffic flow model based on the aerodynamic method to describe the driving behavior of vehicle queues and analyze the influence of intelligent vehicle permeability. Levin M W et al. [44] developed a vehicle following model for predicting scene capacity and wave speed in combination with driver response time. The research results show that travel time has a linear relationship

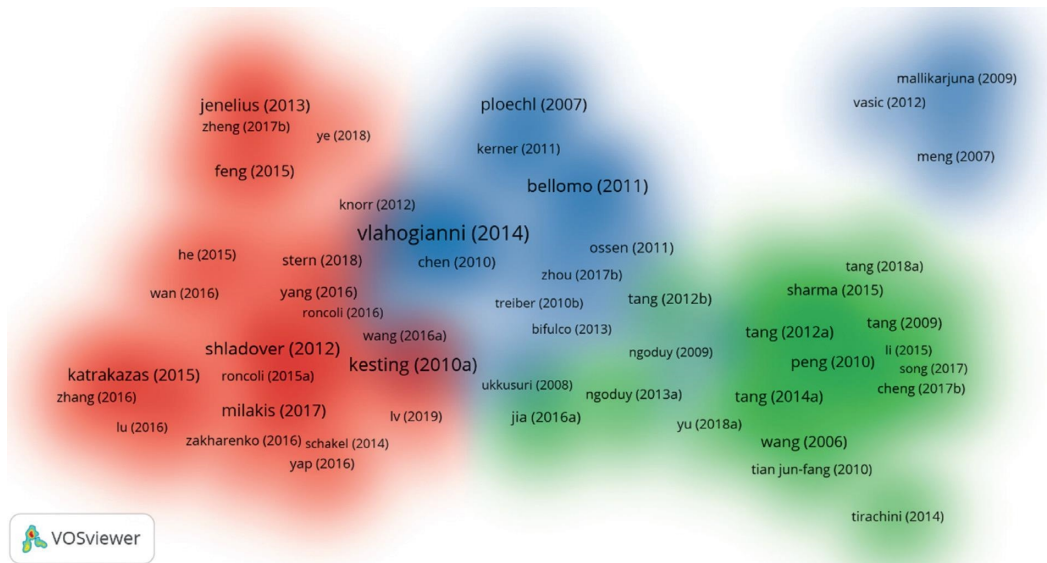


FIGURE 5: Cluster density view of document cocitation network.

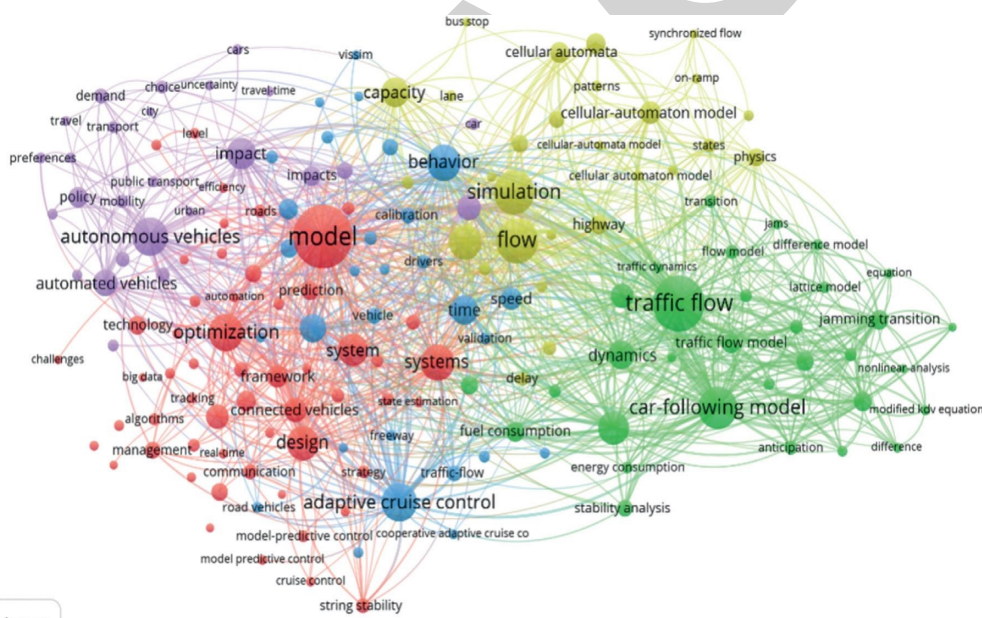


FIGURE 6: Keywords cooccurrence network of intelligent vehicles studies.

with the proportion of autonomous vehicles, and the travel time will be significantly reduced when the penetration rate of autonomous vehicles reaches over 80%.

**3.7.2. Group 2 (Green): Analysis of Macroscopic Traffic Flow Characteristics of Heterogeneous Traffic Flow.** This cluster contains a total of 32 keywords, mainly including traffic flow, stability, waves, fuel consumption, and jamming transition.

In traffic flow research, the impact of autonomous vehicles on traffic flow stability and traffic flow capacity is an important research topic. Many current works have explored the traffic capacity and stability of heterogeneous traffic flow

and intelligent traffic. Kesting et al. [44] proposed an intelligent driver model, that is, automatic real-time detection of traffic conditions based on local information, to adapt to the driving characteristics of ACC.

The results of the test show that at a 25% penetration rate, ACC vehicles can eliminate congestion during specific peak periods of the simulation. Shladover et al. [24] conducted a simulation study on the traffic capacity of single-lane expressways with CACC vehicles mixed with different permeability, and the results showed that the ACC model could not significantly improve the traffic capacity, while the CACC model with medium and high permeability could greatly improve the traffic capacity. Arnaout et al. [45]

studied the influence of permeability on traffic capacity under different traffic demand conditions and found that CACC permeability had a significant impact on traffic capacity when traffic demand was high. At the same time, Arnaout et al. [46] studied the impact of different CACC permeability on traffic efficiency under different traffic demands, and the results showed that CACC permeability had no impact on traffic capacity when traffic demand was low. However, when the traffic demand is high and the permeability exceeds 40%, the permeability has a substantial impact on the increase of traffic capacity.

*3.7.3. Group 3 (Blue): Analysis of Driving Behavior in Heterogeneous Traffic Flow.* There are 29 keywords in this cluster, mainly including behavior, time, adaptive cruise control, safety, risk, and accidents.

When a car is running, the driver needs to control the car based on information provided by the environment, road signs, signals, and instrumentation in the car. Information can be obtained through visual, tactile, and auditory means. Among them, more than 80% of the driver's information is obtained through visual channels, followed by hearing. Driving behavior can be seen as a repetitive message processing model composed of information perception, judgment, and decision-making and actions; that is, judgment and decision-making are the execution of manipulation actions after receiving the perception signal [47]. In the traditional traffic flow field, the analysis of driving behavior generally adopts subjective questionnaire data and objective driving data. Through a reasonable questionnaire, drivers can be divided into various types (such as aggressive, cautious, slow). However, drivers tend to hide their usual aggressive driving behaviors when conducting self-evaluation through the questionnaire, so the collected questionnaire cannot truly describe drivers' driving behaviors. The collected objective driving data of drivers can be used to classify drivers according to clustering analysis or principal component analysis, but during data collection, drivers will feel that they are in a monitoring environment and change their real driving behavior.

In the field of heterogeneous traffic flow, researchers have built different driver models mainly to stimulate and recognize the behavior of human-driven vehicles in heterogeneous traffic flow. These researches are mainly divided into driver behavior models based on control theory and driver behavior models based on machine learning theory. The driver behavior model based on control theory mainly assumes that the driving trajectory is known, studies the driver and the vehicle as a whole, and predicts the vehicle trajectory in the future in real-time according to the historical running state of the vehicle. For example, Song et al. [48] established a "single-lane auto-manual driving mixed traffic flow model considering driving behavior" and a "three-lane auto-manual driving mixed traffic flow model considering driving behavior" based on cellular automata and conducted simulation tests. Zhu et al. [49] used Bando's model and a modified Bando's model to describe human driver behavior and autonomous vehicle, respectively, and

conducted stability analysis on heterogeneous traffic flow. Zheng et al. [50] established a "random model of auto-manual driving mixed traffic flow" based on fully considering the uncertainty of human driving behavior and the interaction between automatic driving and manual driving vehicles. The driver behavior model based on machine learning theory uses the big data of vehicle operation to obtain the characteristic information of the driver and then describes the decision-making process of the driver, to establish the driver behavior model. Wang et al. [51] organized drivers to carry out the following test, analyzed the difference of speed difference cycle in the following process, divided driver types according to driving parameters through cluster analysis, and established an improved IDM model. Michael et al. [44] developed the following model including driver reaction time to study the traffic capacity and backward wave speed of autonomous vehicles under different permeability. Chen et al. [52] used the driver's driving behavior data obtained by the simulation platform to classify the characteristics of the following behavior obtained and solved the problem of operation difference caused by sequence transformation of vehicles in the fleet.

*3.7.4. Group 4 (Yellow): Simulation Experiment Analysis of Heterogeneous Traffic Flow.* There are 28 keywords in this cluster, mainly including simulation, cellular-automaton model, delay, and states.

With the gradual maturity of autonomous driving technology, the heterogeneous traffic flow phenomenon formed by autonomous driving vehicles and manual driving vehicles must exist for a long time. Due to the uncertainty of heterogeneous traffic flow, such as safety, coordination ability, and efficiency, people have raised doubts about the safety of autonomous vehicles. Although the above problems can be verified by establishing a real road test field for real vehicle test, the cost of real vehicle test is too high, there are too few test bases with test qualification, and the test results are susceptible to the influence of the environment and test conditions. The traffic flow simulation test has the characteristics of repeatability, safety, economy, and effectiveness, which can facilitate researchers to carry out tests at lower costs under the same initial conditions.

A simulation model is a mathematical/logical representation of a real-world system, which takes the form of software that is executed experimentally on a digital computer. Common simulation models used in traffic flow include CORSIM, MITSIM, and VISSIM, but most of these models are based on homogeneous traffic conditions and are not suitable for studying the characteristics of heterogeneous traffic flow. Therefore, Vander Werf et al. [53] used Monte Carlo simulation to estimate the impact of vehicle types on lane capacity based on models in the literature and conducted a sensitivity analysis of ACC and CACC vehicle time-interval parameters in simulation tests. Van Arem et al. [54] applied the intelligent vehicle test simulation platform to simulate and analyze the influence of different CACC vehicle ratios on traffic capacity because of the bottleneck section of the expressway with fewer lanes. Talebpour and Mahmassani

[55] applied the ACC vehicle and CACC vehicle following model and conducted numerical simulation experiments to present scatter diagrams of mixed traffic flow concerning density under different traffic demands and different CACC/ACC vehicle ratios. The influence of CACC vehicles and ACC vehicles on-ramp capacity is analyzed. Based on simulation experiments, Lioris [56] studied the impact of the Internet of vehicles' fleet on the capacity of urban road system intersections. Hartmann et al. [57]. studied the influence of CAV on the capacity of the German expressway network through a micro traffic flow simulation experiment. Bujanovic and Lochrane [58] quantitatively analyzed the influence of CAV on the capacity of basic sections of expressways and applied VISSIM simulation software to verify and analyze the proposed model.

*3.7.5. Group 5 (Purple): Policies and Barriers in Heterogeneous Traffic Flow.* There are 26 keywords in this cluster, mainly including policy, impacts, delay, and emissions.

It is widely believed that autonomous driving can provide safer and more comfortable driving and is an important part of urban intelligent transportation systems. In this regard, the United States, Japan, Europe, and other developed countries and regions have increased their research and development and investment in autonomous driving in terms of policies. In particular, the United States has already carried out strategic planning and technical testing of autonomous driving technology as an important part of intelligent transportation systems at the national level.

However, on the other hand, the application of autonomous vehicles may also have adverse effects on the existing traffic system, such as interfering with road traffic, affecting the safety of other vehicles, and other potential threats. Due to the high initial cost of autonomous vehicles, unclear accident liability details, and uncertain privacy issues, the acceptance of autonomous vehicles is not high among the general public. Therefore, many domestic and foreign scholars have a strong interest in the policies and obstacles of autonomous driving vehicles.

Daniel [59] explores the feasible aspects of AVs and discusses their potential impacts on the transportation system. And this paper explores the feasible aspects of AVs and discusses their potential impacts on the transportation system. Another paper "Policy and Society Related Implications of Automated Driving: A Review of Article and Directions for Future Research" by Milonakis is published in the *Journal of Intelligent Transportation Systems*. In this paper, the potential effects of automated driving that are relevant to policy and society are explored, findings discussed in the article about those effects are reviewed, and areas for future research are identified [60].

*3.8. Analysis of Burst Detection: Research Trends of Heterogeneous Traffic Flow.* The burst detection algorithm was first proposed by Kleinberg in 2003. It identifies keywords with high-frequency density characteristics in the article by detecting the density of keyword frequency changes. Burst detection analysis can be used to detect the

frequency of sudden increase of words in topics and keywords and to obtain the start time, end time, and weight of sudden keywords. Through this information, the emergent hot spots and research trends in the field of heterogeneous traffic flow research can be analyzed. CiteSpace regards this kind of mutation information as a way to measure deeper changes. Burst detection in CiteSpace can be used for two types of variables: one is the frequency of words or phrases used in the cited documents, and the other is the frequency of citations obtained from the cited documents. Import the data into the CiteSpace tool and run it for burst detection. The parameters are set as follows:  $\gamma=1.0$ , the number of states = 2.0, and minimum duration = 2. Select the top 20 meaningful keywords with the largest burst weight for visual analysis, as shown in Figure 7.

As shown in Figure 7, "Start year" and "end year" represent the year of burst keyword duration from start to end. Intensity represents the frequency of keyword occurrences. From the point of view of burst intensity, cellular automata, traffic flow, numerical simulation, and other links have high intensity. It can be seen that traditional traffic flow studies focus on using the cellular automata model for modeling and carry out numerical simulation calculation; related articles include spatial-temporal patterns in heterogeneous traffic flow with a variety of driver behavioral characteristics and vehicle parameters. Different from the traditional heterogeneous traffic flow research, heterogeneous traffic flow research in the new period is mostly dependent on artificial intelligence, cloud computing, big data, and other new-generation network technologies represented by the Internet of vehicles. Related articles include Linear Stability Analysis of Heterogeneous Traffic Flow considering Degradations of Connected Automated Vehicles and Reaction Time. Capacity and stability are macro characteristics of heterogeneous traffic flow, which have a great impact on the safety and efficiency of heterogeneous traffic flow. Related articles include Controllability Analysis and Optimal Control of Mixed Traffic Flow with Human-Driven and Autonomous "Vehicles," Stability Analysis, and the Fundamental Diagram for Mixed Connected Automated and Human-driven Vehicles and Analysis on Traffic Stability and Capacity for Mixed Traffic Flow with Platoons of Intelligent Connected Vehicles.

The burst intensity of the dynamic model, car-following model, and adaptive cruise control was 29,108, 55,438, and 3085, respectively. Adaptive cruise control not only has the function of cruise control but also adaptively adjusts the motion state of the vehicle in front according to the motion state of the vehicle, which has a certain influence on the traffic flow research: "Cooperative Adaptive Cruise Control and Exhaust Emission Evaluation Under Heterogeneous Connected Vehicle Network Environment in Urban City," "Cooperative Adaptive Cruise Control and Intelligent Traffic Signal Interaction: A Field Operational Test with Platooning on a Suburban Arterial in Real Traffic," and "Longitudinal Emissions Evaluation of Mixed (Cooperative) Adaptive Cruise Control Traffic Flow and Its Relationship with Stability."



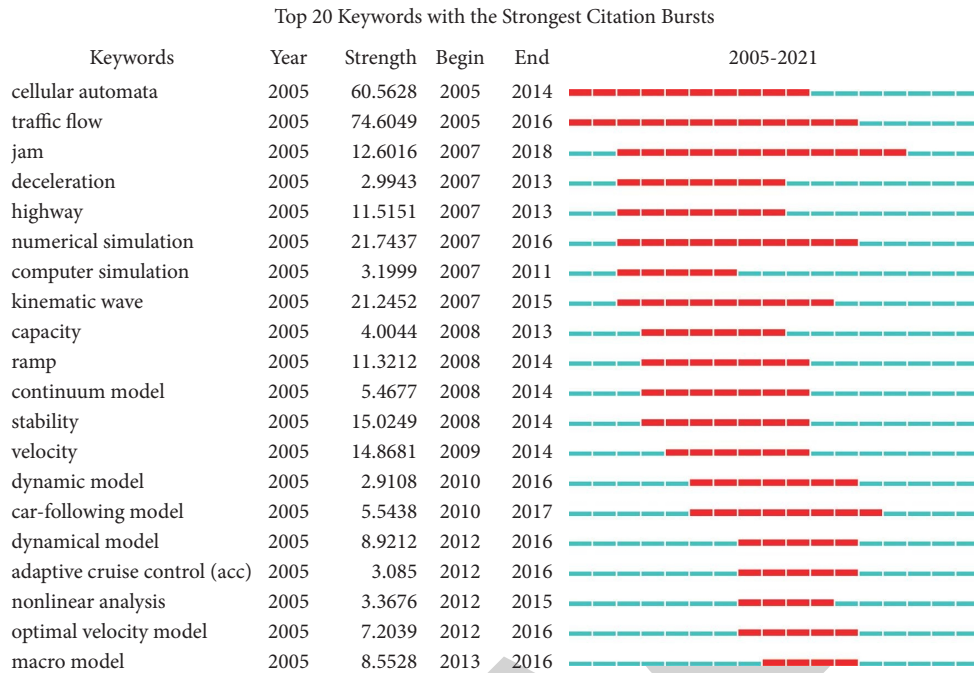


FIGURE 7: Keywords of the top 20 burst weights from 2004 to 2021.

#### 4. Conclusion

This paper uses VOSviewer software to conduct a bibliometric analysis of 4709 documents related to heterogeneous traffic flow, from the annual distribution statistics of the article, national and regional statistics, analysis of major research institutions, analysis of coauthors of major research groups, analysis of source journals, and key. In terms of word and burst detection analysis, the research results in the field of heterogeneous traffic flow in recent years (2004–2021) are analyzed from multiple angles, and the existing article review in this field is supplemented to consolidate the research results in the field of heterogeneous traffic flow. The analysis results show that cocitation analysis divides heterogeneous traffic flow into three main research directions, which include “heterogeneous traffic flow model,” “traffic flow capacity analysis,” and “traffic flow stability analysis.” The keyword cooccurrence analysis is applied to identify five dominant clusters: “modeling and optimization methods,” “traffic flow characteristics analysis,” “driving behavior analysis,” “simulation experiment,” and “policies and barriers.” The main research conclusions are as follows:

- (1) The number of documents in the field of heterogeneous traffic flow research continues to grow, indicating that the international academic community pays more attention to the field of heterogeneous traffic flow, and related research is also in progress. According to the distribution of the article in various countries, China, the United States, and India rank high, indicating that these countries are active areas for heterogeneous traffic flow research. In terms of research institutions, Beijing Jiaotong

University, Southeast University, and Tongji University have the highest research output, and the cooperation between major research institutions has been further strengthened. Judging from the analysis results of the main author groups, the international exchanges and cooperation among highly productive authors are quite active. In terms of source journals, “transportation research part c-emerging technologies,” “physical a-statistical mechanics and its applications,” and “transportation research record” are authoritative journals in the field of heterogeneous traffic flow research and are also important platforms for publishing research results.

- (2) Through the analysis of the cooccurrence network diagram of heterogeneous traffic flow keywords, the research directions represented by these keywords can be divided into four categories: “modeling and optimization methods,” “macroscopic traffic flow characteristics analysis,” “driving behavior analysis,” “simulation experiment analysis,” and “policies and barriers.” These five representative keywords also reflect the research trends in this field in recent years. The establishment of a safe and reliable model is of great significance to the development of autonomous driving and has received more and more attention in recent years. Related researches on the macroscopic traffic flow characteristics of heterogeneous traffic flow are also emerging. In addition, with the continuous development of information technology, sensor technology, computer technology, and other technologies, there are also many studies exploring the improvement of people’s living standards and traffic environment from the perspective of



simulation from the perspective of heterogeneous traffic flow.

- (3) This study also has some shortcomings. The search keywords used in this research are representative of the research in this field and cannot completely contain all the keywords in this research field. In addition, the search scope of this study is only for the SCIE and SSCI citation index databases in the core collections of the Web of Science (WOS). This study does not discuss the proportions and partnerships of each topic. At the same time, this research aims to provide a visual analysis method using the MKD method, which is also applicable to other keywords and databases. In terms of visualization methods, different goals based on analysis of different visualization methods can be further considered in future work.

### Data Availability

The datasets used to support the findings of this study are available from the corresponding author upon request.

### Conflicts of Interest

The authors declare that they have no conflicts of interest.

### Acknowledgments

This study was sponsored by the National Key Research and Development Project (2020YFE0201200) and the National Natural Science Foundation of China (52072292).

### References

- [1] Z. Yao, B. Zhao, T. Yuan, H. Jiang, and Y. Jiang, "Reducing gasoline consumption in mixed connected automated vehicles environment: a joint optimization framework for traffic signals and vehicle trajectory," *Journal of Cleaner Production*, vol. 265, Article ID 121836, 2020.
- [2] L. Roger, "Accelerating the future: the economic impact of the emerging passenger economy," *Strategy Analytics*, 2017.
- [3] Q. Wang, *The Road from the Internet of Vehicles to the Connected and Intelligent Traffic Network of Autonomous Vehicles*, Posts and Telecommunications Press, Beijing, China, 1st edition, 2018.
- [4] Z. Xin and F. David, *Review of International Oil Consumption Trends and Policies*, Lawrence Berkeley National Laboratory, Berkeley, CA, USA, 2019.
- [5] T. Ersal, I. Kolmanovsky, N. Masoud et al., "Connected and automated road vehicles: state of the art and future challenges," *Vehicle System Dynamics*, vol. 58, no. 5, pp. 672–704, 2020.
- [6] S. Liu, "The first unmanned driving technology book," Publishing House of Electronics Industry, Beijing, China, 1st edition, 2017.
- [7] L. Yang, X. Zhao, and G. Wu, "A review of cooperative eco-driving strategies for intelligent networked Vehicles," *Journal of Traffic and Transportation Engineering*, vol. 20, no. 5, pp. 58–72, 2020.
- [8] J. Liu, N. Wu, Y. Qiao, and Z. Li, "A scientometric review of research on traffic forecasting in transportation," *IET Intelligent Transport Systems*, vol. 15, no. 1, pp. 1–16, 2021.
- [9] F. Guo, W. Lv, L. Liu, T. Wang, and V. G. Duffy, "Bibliometric analysis of simulated driving research from 1997 to 2016," *Traffic Injury Prevention*, vol. 20, no. 1, pp. 64–71, 2019.
- [10] H. Small, "Co-citation in the scientific article: a new measure of the relationship between two documents," *Journal of the American Society for Information Science and Technology*, vol. 2, no. 24, pp. 265–269, 1973.
- [11] K. C. Garg and H. K. Tripathi, "Bibliometrics and scientometrics in India: an overview of studies during 1995–2014 Part I: Indian publication output and its citation impact," *Annals of Library and Information Studies*, no. 64, pp. 28–36, 2017.
- [12] H. D. White and K. W. McCain, "Visualizing a discipline: an author co-citation analysis of information science, 1972–1995," *Journal of the American Society for Information Science*, vol. 49, no. 4, pp. 327–355, 1998.
- [13] B. Wang, "Ten years research and development of bibliometrics in China (1979–1989)," *Information Science*, vol. 13, no. 1, pp. 56–64, 1992.
- [14] Y. He, S. Yang, C. Chan, and L. Chen, "Visualization analysis of intelligent vehicles research field based on mapping knowledge domain," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, pp. 1–16, 2020.
- [15] X. Zou, W. L. Yue, and H. L. Vu, "Visualization and analysis of mapping knowledge domain of road safety studies," *Accident Analysis & Prevention*, vol. 118, pp. 131–145, 2018.
- [16] X. Zou and H. L. Vu, "Mapping the knowledge domain of road safety studies: a scientometric analysis," *Accident Analysis & Prevention*, vol. 132, Article ID 105243, 2019.
- [17] F. Lu and L. Fuhai, "Advances in co-word analysis," *Journal of Library Science in China*, no. 2, pp. 88–92, 2006.
- [18] X. Zhang, X. Wang, and Y. Ma, "International research progress on driving behavior and driving risk," *China Journal of Highway and Transport*, vol. 33, no. 6, pp. 1–17, 2020.
- [19] X. An, X. Huang, and X. Zhang, "Analysis of the key words of journal work bibliometrics," *Chinese Journal of Scientific and Technological Journals*, vol. 13, no. 6, pp. 505–506, 2002.
- [20] C. Chen, "CiteSpace II: detecting and visualizing emerging trends and transient patterns in scientific literature," *Journal of the American Society for Information Science and Technology*, vol. 57, no. 3, pp. 359–377, 2006.
- [21] N. J. van Eck and L. Waltman, "Software survey: VOSviewer, a computer program for bibliometric mapping," *Scientometrics*, vol. 84, no. 2, pp. 523–538, 2010.
- [22] N. J. van Eck, L. Waltman, and J. van den Berg, "Visualizing the WCCI 2006 Knowledge Domain," in *Proceedings of the 2006 IEEE International Conference on Fuzzy Systems*, IEEE, Vancouver, Canada, July 2006.
- [23] A. Kesting, M. Treiber, and D. Helbing, "Enhanced intelligent driver model to access the impact of driving strategies on traffic capacity," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 368, no. 1928, pp. 4585–4605, 2010.
- [24] S. E. Shladover, D. Su, and X.-Y. Lu, "Impacts of cooperative adaptive cruise control on freeway traffic flow," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2324, no. 1, pp. 63–70, 2012.
- [25] E. I. Vlahogianni, M. G. Karlaftis, and J. C. Golias, "Short-term traffic forecasting: where we are and where we're going," *Transportation Research Part C: Emerging Technologies*, vol. 43, pp. 3–19, 2014.
- [26] X. Chen, L. Li, and Y. Zhang, "A Markov model for headway/spacing distribution of road traffic," *IEEE Transactions on Intelligent Transportation Systems*, vol. 11, no. 4, pp. 773–785, 2010.

## Research Article

# An Improved Simulated Annealing-Based Decision Model for the Hybrid Flow Shop Scheduling of Aviation Ordnance Handling

Xianglei Meng , Nengjian Wang , Jue Liu , and Qinhui Liu 

College of Mechanical and Electrical Engineering, Harbin Engineering University, Harbi 150001, China

Correspondence should be addressed to Qinhui Liu; [liuqinhui@hrbeu.edu.cn](mailto:liuqinhui@hrbeu.edu.cn)

Received 20 December 2021; Revised 12 January 2022; Accepted 15 January 2022; Published 2 February 2022

Academic Editor: Daqing Gong

Copyright © 2022 Xianglei Meng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Aviation ordnance handling is critical to the firepower projection of the time-critical cyclic flight operation on aircraft carriers. The complexity of the problem depends on the supply and demand features of ordnance. This paper examines the scheduling of aviation ordnance handling of an operational aircraft carrier under the framework of hybrid flow shop scheduling (HFS) and derives a method based on the simulated annealing (SA) algorithm to get the HFS problem's solution. The proposed method achieves the minimum possible flow time by optimizing the ordnance assignment through different stages. The traditional SA algorithm depends heavily on the heuristic scheme and consumes too much time to compute the optimal solution. To solve the problem, this paper improves the SA by embedding a task-based encoding method and a matrix perturbation method. The improved SA remains independent of the heuristic scheme and effectively propagates the local search process. Since the performance of SA is also influenced by its embedded parameters, orthogonal tests were carried out to carefully compare and select these parameters. Finally, different ordnance loading plans were simulated to reveal the advantage of the improved SA. The simulation results show that the improved SA (ISA) can generate better and faster solution than the traditional SA. This research provides a practical solution to stochastic HFS problems.

## 1. Introduction

This study focuses on the ordnance dispatching scheduling problem observed onboard the aircraft carrier flight operation, which plays an important role in the air wing firepower projection in its sortie generation [1]. The ordnance handling process involves many stages, equipment, and hundreds of personnel operating in a limited work space [2]; finding an optimal dispatch scheduling for a given ordnance load plan, plus the time critical nature of cyclic flight window requires, is a challenging problem. Traditionally, ordnance dispatching scheduling is made by a human operator's hand in a spreadsheet with experience, which is always nonoptimal, or even leading to delays that left aircraft launching without firepower. Thus, robust optimal scheduling is essential to conduct the ordnance handling procedure. However, such problem has seldomly been studied, which can be casted in the hybrid flow shop (HFS) scheduling framework.

The HFS scheduling problem [3] can be regarded as the combination of the flow shop scheduling (FSS) [4] problem

and parallel machine scheduling (PMS) problem, where the former is to decide the job sequences through the shop and the latter is to allocate jobs to machines, given the processing times of each job on each machine according to one or several given criteria, aiming to minimize the makespan [5–7]. For an  $n$  jobs  $m$  stages problem, there are a total of  $(n!)^m$  possible schedules, which proves to be NP-hard [8]. If the numbers  $n$  of jobs and  $m$  of stages are very small, the optimal schedule may be determined by exhaustion, such as branch and bound (B&B) [9] or integer programming techniques [10]. However, these approaches are not applicable to HFS problems with numerous jobs and stages, due to their enormous computing time and memory occupation.

Thus, for scheduling different HFS configurations, a large number of approximation and heuristics algorithms have been proposed [11, 12]. The computational complexity of HFS propelled scholars to develop many heuristics algorithms to obtain good enough solutions in a short time for medium-to-large problems, such as different scheduling rules [13], but the heuristic methods are too problem-

specific, it often cannot be applied to generalized problem. For the past decades, many general schemes on improving the performance of simple heuristics have been successfully developed, most of which are named as metaheuristics, such as genetic algorithm (GA) [14], ant colony optimization (ACO) [15], tabu search (TS) [16], neural networks (NN) [17], artificial immune systems (AIS) [18], and simulated annealing (SA) [19]. They inhere with higher level of abilities in searching the vast solution space, which have better performance than the simple heuristic methods.

Since different heuristics work effectively for different problems, when it encounters the flow shop scheduling problem, Maaroju [20] tested all the metaheuristic methods and found that the genetic algorithm and simulated annealing outperformed others, for hill climbing, swarm intelligence, and neural networks yielded only marginal improvements. However, the computation time for the GA is larger than that for SA. Thus, the SA-based algorithm is chosen in dealing with the ordnance handling problem under the HFS framework. Simulated annealing origins from the metallurgy technology, where a material cools down from high temperature to get minimum energy state. In the algorithm, the current state  $s$  and neighbor states  $s'$  are considered, and the algorithm decides the state transition probability from  $s$  to  $s'$  based on current system energy (known as temperature). This process continues until a good enough state has been found or the computation threshold has been reached. Such mechanism guarantees approximating to global optimum without getting stuck in local minimum for solving large complex optimization problems. However, the traditional SA algorithm has several defects [21], which include heuristic-dependent, parameter-specific, and long computation time; thus, the performance of the algorithm is yet to be improved. To overcome the above defects, this paper presents a Monte Carlo [22] perturbation method, which directly perturbs the solution matrix in each iteration of SA cooling, eliminating the dependence on any heuristic method, whereas SA performance also depends on cooling parameters; this paper carefully plans the calibration of these parameters to accelerate the computation process by adding double thresholds and setting the memory method of the SA.

The organizations of this paper are as follows: Section 2 introduces the ordnance handling process in detail; Section 3 discusses the methodology of the SA and improves the SA by embedding a new decoding method and matrix perturbation method; Section 4 evaluates the improved SA algorithm through computational experiments; Section 5 summarizes the research findings and gives the directions of future research.

## 2. Ordnance Handling Procedure

The ordnance handling procedure is specified in a daily loading plan, which lists the amount and types of weapons (throughout this paper, ordnance and weapon are used interchangeably) to be loaded onto the corresponding aircraft. Figure 1 shows the layout of aircraft carrier decks, where the construction and transfer of ordnance origin from the magazines located in lower decks to the awaiting aircraft

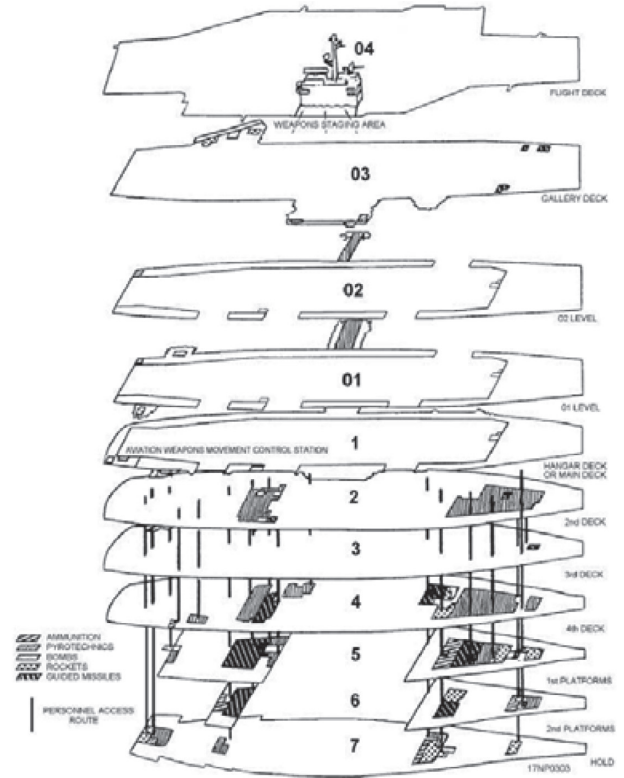


FIGURE 1: The layout of the ordnance handling routes aboard an aircraft carrier.

on flight deck following a series of stages. In stage I, the ordnances are retrieved from magazines by bomb skids and delivered to lower-stage elevators. In stage II, the ordnances are lifted to the hangar deck by lower-stage elevators. In stage III, the ordnances are transferred to the staging area of the hangar deck, assembled in that area, and moved to upper-stage elevators. In stage IV, the ordnances are transferred to the flight deck by upper-stage elevators. In stage V, the ordnances are moved directly to and loaded on the aircraft waiting on the flight deck. The flowchart of this procedure is shown in Figure 2. For a common aircraft carrier layout (Figure 1), there are at least 4 magazines in the delivering stage, 8 elevators in the lower-lifting stage, 2 assemble centers in the assembling stage, 4 elevators in the upper-lifting stage, and around 10 aircraft spots in the loading stage. According to Gupta [8], the two-stage flow shop problem with one stage containing a single machine can be NP-hard. Thus, the ordnance handling problem is far from trivial, especially for ordnance officers in making timely decisions of the flight operations.

This paper examines the ordnance handling problem in the HFS framework. The definition of hybrid flow shop system is as follows: in a factory, the set of  $n$  jobs  $J = \{1, 2, \dots, j, \dots, n\}$  is going to be processed through  $m$  stages  $M = \{1, 2, \dots, i, \dots, m\}$  in sequence, while each stage  $i$  contains  $M_i = \{1, 2, \dots, k, \dots, m_i\}$  identical machines, and the processing time of job  $j$  on machine  $k$  is  $p_{jk} \geq 0$ . The objective is always to minimize makespan. Similarly, in the ordnance handling problem,  $n$  batches of weapons are

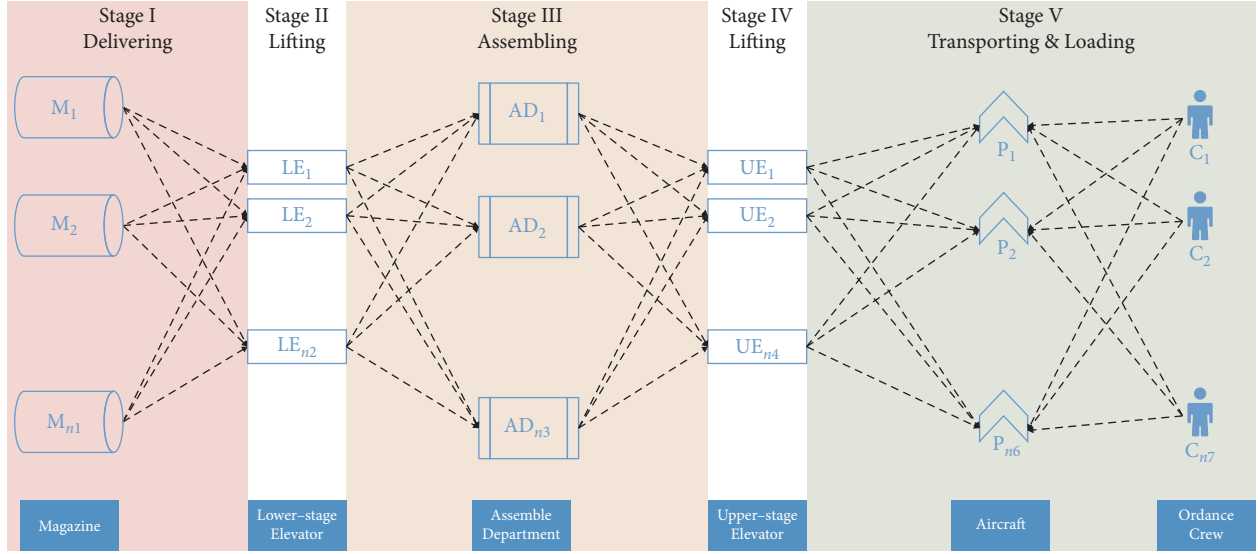


FIGURE 2: Flowchart of ordnance handling process. Note:  $M$  = magazine;  $LE$  = lower-stage elevator;  $AD$  = assembly department;  $UE$  = upper-stage elevator;  $P$  = aircraft parking spot.

considered as jobs; they also shall be processed in the same order through  $m$  stages by facility  $k$  (as shown in Figure 2) with processing time  $p_{jk}$ , and the objective is to decide the weapons' sequences and the allocations of weapons to facilities to get the minimum flow time. This is a combinatorial optimization problem with  $(n!)^m$  possible schedules, which is considered as NP-hard so that it is difficult to find the optimal solution in polynomial time. For a simple case of 10 batches of weapons in our problem, there can be  $(10!)^5 = 6.3 \times 10^{32}$  different schedules for

the ordnance officer to choose, which is beyond human mind's reach in conducting the time-critical flight operations.

The ordnance configuration of aircraft on the carrier depends on the specific mission [23]. It is assumed that each transfer equipment or personnel only transfers one type of ordnance at a time [24]. For each type of ordnance, the number loaded in one skid is denoted as  $r_w$ . Thus, the batches of ordnances needed to complete the task of all aircraft can be determined by

$$\text{task} = \left\{ \begin{array}{l} 1 \ (1, \text{num}_1) \ (2, \text{num}_2) \ \dots \ (w, \text{num}_w) \\ 2 \ (1, \text{num}_1) \ (2, \text{num}_2) \ \dots \ (w, \text{num}_w) \ \dots \ (W, \text{num}_W) \\ \dots \\ a \ (1, \text{num}_1) \ (2, \text{num}_2) \ \dots \ (w, \text{num}_w) \ \dots \ (W, \text{num}_W) \end{array} \right\}, \quad w = 1, 2, \dots, W, \quad (1)$$

where  $a$  is the number of aircraft to be loaded,  $w$  is the type of ordnance, and  $\text{num}_w$  is the actual number of ordnance types  $w$ .

$$\xi = \sum_{i=1}^a \sum_{j=1}^w \frac{\text{num}_w}{r_w}, \quad i = 1, 2, \dots, a, \quad j = 1, 2, \dots, w, \quad (2)$$

where  $r_w$  is the number of type  $w$  ordnances in one skid,  $a$  is the number of aircraft, and  $w$  is the number of ordnance types to be loaded on the aircraft. That is, an ordnance should be transferred to the required aircraft, once being retrieved from the magazine.

**2.1. Stage I: Weapons Retrieving.** Multiple magazines are located in the bow and aft of the carrier, and the ordnances can be transferred by multiple elevators. The ordnances are firstly retrieved from magazines by skids with the setup time

$T_0$ . From the same magazine, the ordnances should be retrieved with an interval no shorter than  $t_{\text{int}}$ . Then, the skids deliver the ordnances to lower-stage elevators. The time consumed to transfer ordnances from magazines to lower-stage elevators can be expressed as

$$T_M^L = \begin{bmatrix} t_1^1 & t_1^2 & \dots & t_1^L \\ t_2^1 & t_2^2 & \dots & t_2^L \\ \vdots & \vdots & \ddots & \vdots \\ t_M^1 & t_M^2 & \dots & t_M^L \end{bmatrix}, \quad (3)$$

where  $M$  is a magazine and  $L$  is a lower-stage elevator.

**2.2. Stage II: Weapons Buildup.** The ordnances are loaded onto lower-stage elevators and lifted vertically with constant speed to the hangar deck. The time consumed in this stage (lifting time) is denoted as  $T_L$ .

2.3. *Stage III: Weapons Assembling.* The ordnances are preassembled in the staging area of the hangar deck, with a sufficient lead time to meet the short turnaround time of the flight schedule. The assembling time  $T_K^{ass}$  varies with the types of ordnances. Note that the assembling time of the staff fluctuates in the real world. Therefore, the interval of assembling time was set to  $[-T_{f1}, T_{f1}]$ . The real assembling time is denoted as  $T_K^{ass} + t_1$ , where  $t_1$  is a random number within  $[-T_{f1}, T_{f1}]$ .

2.4. *Stage IV: Weapons Striking Up.* The ordnances are transferred to the flight deck by upper-stage elevators. The time consumed in this stage (transport time) is denoted as  $T_U$ .

2.5. *Stage V: Weapons Loading.* Some ordnance crew members on the flight deck transport the ordnances from the upper-stage elevators to the aircraft. The time consumed in this stage can be expressed as

$$T_U^A = \begin{bmatrix} t_1^1 & t_1^2 & \cdots & t_1^A \\ t_2^1 & t_2^2 & \cdots & t_2^A \\ \vdots & \vdots & \vdots & \vdots \\ t_U^1 & t_U^2 & \cdots & t_U^A \end{bmatrix}, \quad (4)$$

where  $U$  is an upper-stage elevator and  $A$  is an aircraft.

The other ordnance crew members load the ordnances onto the aircraft. It is assumed that the different groups consume the same time to load the same ordnance and different types of ordnances need different time to be loaded. The time needed to load each type of ordnance is denoted as  $T_K^{load}$ . The interval of the loading time was set as  $[-T_{f2}, T_{f2}]$ . The real loading time is denoted as  $T_K^{load} + t_2$ , where  $t_2$  is a random number within  $[-T_{f2}, T_{f2}]$ .

The ordnance crew members can be shared across groups. The loading cannot proceed unless ordnance crew members are available. The walking time for interstation transfer between different aircraft can be expressed as

$$T_A^A = \begin{bmatrix} t_1^1 & t_1^2 & \cdots & t_1^A \\ t_2^1 & t_2^2 & \cdots & t_2^A \\ \vdots & \vdots & \vdots & \vdots \\ t_A^1 & t_A^2 & \cdots & t_A^A \end{bmatrix}. \quad (5)$$

Referring to the standard three-field notation for scheduling problems, our problem can be described as follows:  $FH5, (PM^{(k)})_k^2 = 1, (RM^{(k)})_{k=3}^4 | prmu, M_j^{(5)}, block | C_{max}$ . Specifically,  $FH5$  is a five-stage HFS problem: stage I involves  $M^{(1)}$  identical magazines that store ordnances; stage II involves  $M^{(2)}$  lower-stage elevators to transport the ordnances; stage III has  $M^{(3)}$  identical assembling personnel to assemble the ordnances; stage IV has  $RM^{(4)}$  independent upper-stage elevators to transport the ordnances; and stage V has  $RM^{(5)}$  independent aircraft to be loaded. Note that  $prmu$  indicates that the ordnances are handled in the same

order in every stage;  $M_j^{(5)}$  (eligibility constraint) means that the handling of ordnance  $j$  is limited to the aircraft set  $M$  in stage V;  $block$  indicates that the capacity of buffer between stages is constrained, for instance, the weapons have to wait in the current stage till enough room is released for the next stage of handling.

In total, the completion time for batch  $i$  of ordnances can be calculated by

$$C_i = T_M^L + T_L + T_K^{ass} + t_1 + T_U + T_U^A + T_A^A + T_K^{load} + t_2 + T_{wating}, \quad i = 1, 2, \dots, \xi, \quad (6)$$

where  $T_{wating}$  is the whole waiting time in the transporting process, for an ordnance cannot be handled unless machines or ordnance crew members are available.

The general objective of ordnance handling is to complete all transporting operations as efficiently as possible within the specified time and to generate a reasonable schedule for ordnance handling. Therefore, for our ordnance handling problem, the minimization of makespan is set as goal, so that sufficient ordnances can be loaded to the awaiting aircraft to fly in the next flying window.

$$\text{Object} = \min(C_{max}). \quad (7)$$

### 3. Improved Simulated Annealing Algorithm

The SA is a technique capable of searching for good solutions to various combinatorial problems in material science and physics. The pseudocode of the algorithm is as follows.

The SA includes three major functions: state generation, state acceptance, and temperature update. The first is to make perturbations of the given initial solution, in order to search for the optimal solution effectively in the vast solution space. The second decides whether to accept a newly generated solution with a certain probability in case of trapping in the local minimum. The third offers a cooling scheme that mimics the physical annealing process to get a stable state for the problem. The SA performance can be augmented by adjusting various parameters and operators [25], such as initial temperature, descent gradient of temperature, and termination rule, which have to be adjusted manually. This paper mainly improves the search efficiency (timeliness) of the SA, without sacrificing the optimization quality. Thus, encoding scheme, initial solution, neighborhood search structure (NSS), type of cooling schedule, and two criteria (internal cycle termination and external cycle termination) are modified as described in the following subsections.

3.1. *Task Encoding.* This paper presents a task-based encoding method for the ordnance handling problem, where ordnances are coupled with aircraft. Each job of the HFS is defined as an operation  $o^{xy}$ , where  $x$  is the type of ordnance and  $y$  is the aircraft to be loaded. Then, a solution can be expressed as



Input: A initial  $S$  solution and a cost value function  $F(x)$ .  
Output: A  $S'$  solution that minimizes the cost value function  $F(x)$ .  
 $T \leftarrow$  initializing Temperature//a method for assigning an initial temperature  
**while**  $T > B$  not freezing **do**//a definition of "frozen,"  
    **for**  $i = 1$  to  $C$  **do** while not at equilibrium **do**//a definition of "equilibrium,"  
         $S' \leftarrow$  new permutation of  $S$ .  
        If  $F(S') < F(S)$  or Random value  $< e^{(F(S') - F(S))/DT}$  then  $S \leftarrow S'$ //a selection criterion  
    **end**  
     $T \leftarrow$  reduced temperature//a way of calculating the next temperature  
**end**

ALGORITHM 1: Standard simulated annealing.

$$S = \begin{pmatrix} o_1^{xy} & m_{11} \cdots & m_{n1} \\ o_2^{xy} & m_{21} \cdots & m_{n1} \\ \vdots & \vdots & \vdots \\ o_i^{xy} & m_{n1} \cdots & m_{nn} \end{pmatrix}, \quad (8)$$

where the first column is a permutation of the task sequence, the following columns are the corresponding stages, and  $m_1$  to  $m_n$  are machines assigned randomly to execute the tasks.

**3.2. Initial Solution.** As mentioned above, the initial solution has a great impact on the final solution of the SA. According to the heuristic methods mentioned in [26], we generate the initial solution by the following rule: first, the permutation of tasks is determined by assigning each type of weapon  $x$  to its corresponding aircraft  $y$ , defined as operation  $o_i^{xy}$ , according to the ordnance loading plan  $P$ . Then, the tasks are assigned to the earliest available machine. If there are more than one earliest available machines, one of them will be chosen randomly. The initial solution will be generated as  $S_{i \times m}$  with  $i$  tasks through  $m$  stages.

For example, Table 1 shows the ordnance loading plan of the aircraft waiting to operate in the next fly window. There are three aircraft and four types of ordnances. Aircraft 1 requires 2 skids of type 1 ordnances and 1 skid of type 2 ordnances.

Here, the operations are permuted by aircraft number (equal priority) in ascending order:  $o_i^{xy} = \{o_1^{11}, o_2^{11}, o_3^{21}, o_4^{12}, o_5^{32}, o_6^{43}\}$ . First, a random permutation of jobs (tasks) is generated. Then, the optimal available elevator is assigned to each operation. The final plan can be written as

$$\begin{pmatrix} 3 & 1 & 1 & 1 & 1 & 1 \\ 5 & 2 & 2 & 2 & 2 & 2 \\ 1 & 3 & 3 & 1 & 3 & 3 \\ 4 & 4 & 4 & 2 & 4 & 1 \\ 2 & 1 & 1 & 1 & 1 & 2 \\ 6 & 2 & 2 & 2 & 2 & 3 \end{pmatrix}. \quad (9)$$

The meaning of such matrix can be explained as follows: taken the third row for illustration, operation  $o_1^{11}$  is extracted out of the third magazine, transferred by the third lower-stage elevator to the first staging area, moved by the third

upper-stage elevator to the flight deck, and loaded to the corresponding aircraft by the No. 3 ordnance crew.

**3.3. Neighborhood Perturbation.** The NSS can generate a new solution by slightly modifying the current candidate solution. Traditionally, many different NSSs are adopted in each iteration of SA computation, such as swap, shift, and reversion [27], to realize the task permutation in the first stage. Then, the same heuristic rules are applied in the following stages to generate a new solution  $S'$ . The quality of the solution depends on the selected heuristic rules. If the rules are too greedy, the algorithm may fall into the local optimum, being unable to get convergence to the final optimal result.

The quality of the SA solution is highly sensitive to the selection of candidate solutions. Therefore, the perturbation scheme is crucial to the good performance of the SA algorithm. To ease the dependence of neighborhood search on heuristics, this paper proposes a Monte Carlo perturbation technique, which directly perturbs the initial solution matrix. The initial solution is changed significantly in one step, eliminating the effect by heuristic methods. The matrix perturbation is described as follows.

The following is an example of the matrix perturbation process: for instance, a  $6 \times 3$  matrix, a rectangle  $R_{a \times b}$  randomly generated a size of  $4 \times 3$  matrix.

$$\begin{pmatrix} 3 & 1 & 2 \\ 4 & 2 & 1 \\ 6 & 3 & 2 \\ 2 & 1 & 3 \\ 1 & 2 & 1 \\ 5 & 3 & 3 \end{pmatrix} \text{ which covers the matrix of } \begin{pmatrix} 4 & 2 & 1 \\ 6 & 3 & 2 \\ 2 & 1 & 3 \\ 1 & 2 & 1 \end{pmatrix}.$$

$$\text{Then, matrix } R \text{ is reversed as } R' = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 3 \\ 6 & 3 & 2 \\ 4 & 2 & 1 \end{pmatrix}.$$

$$\text{Then, the new solution can be obtained as } S' = \begin{pmatrix} 3 & 1 & 2 \\ 1 & 2 & 1 \\ 2 & 1 & 3 \\ 6 & 3 & 2 \\ 4 & 2 & 1 \\ 5 & 3 & 3 \end{pmatrix}.$$



TABLE 1: Ordnance loading plan.

Aircraft type of ordnance	1	2	3
1	2	1	—
2	1	—	—
3	—	1	—
4	—	—	1

- (1) Randomly generate a rectangle  $R_{a \times b}$ , with width  $a \in (0, i]$  and length  $b \in (0, m]$ ;
- (2) Put  $R$  in the solution matrix  $S$  and mark the four vertices  $(o_i^{xy}, o_{i+a}^{xy}, o_{i+a}^{x+b, y+b}, o_i^{x+b, y+b})$ ;
- (3) The subsequence of the  $R$  area is reversed,  $o_i^{xy} \rightarrow o_{i+a}^{x, y}$  and  $o_i^{x+b, y+b} \rightarrow o_{i+a}^{x+b, y+b}$ ;
- (4) Get the new solution  $S' = \begin{pmatrix} o_{i+a}^{xy} & \dots & o_{i+a}^{x+b, y+b} \\ \vdots & \ddots & \vdots \\ o_i^{xy} & \dots & o_i^{x+b, y+b} \end{pmatrix}$ .

ALGORITHM 2: Matrix perturbation.

Hence, the matrix perturbation process is completed:

$$\begin{pmatrix} 3 & 1 & 2 \\ 4 & 2 & 1 \\ 6 & 3 & 2 \\ 2 & 1 & 3 \\ 1 & 2 & 1 \\ 5 & 3 & 3 \end{pmatrix} \rightarrow \begin{pmatrix} 3 & 1 & 2 \\ 1 & 2 & 1 \\ 2 & 1 & 3 \\ 6 & 3 & 2 \\ 4 & 2 & 1 \\ 5 & 3 & 3 \end{pmatrix}.$$

**3.4. Cooling Schedule.** The SA behavior can be regulated by the temperature and its descent gradient. To avoid the local optimum trap, inferior solutions may be accepted depending on the falling temperature, under the mechanism of cooling schedule. Here, the exponential cooling rate is adopted:

$$T_l = \frac{(T_0 - T_f)(N + 1)}{N(l + 1)} + T_0 - \frac{(T_0 - T_f)(N + 1)}{N}; \quad l = 1, 2, \dots, N, \quad (10)$$

where  $T_0$  and  $T_f$  are initial temperature and final temperature, respectively, and  $N$  is the number of temperatures between  $T_0$  and  $T_f$ .

The SA needs to accept the new state through probability judgment, in order to avoid the local minimum. When the initial temperature is sufficiently high, the cooling is slow enough (i.e., each temperature is held for a sufficiently long time), and the final temperature approaches zero; the SA will converge to the global optimal solution with the probability of 1. However, it is very difficult to fulfil the global convergence condition. Besides, the current state may be worse than some intermediate states in the search trajectory, owing to the probability acceptance mechanism. Thus, the SA algorithm often converges to an approximate optimal solution, or a solution poorer than the best intermediate solution. The search efficiency is inevitably affected. To preserve the best-known state and improve search efficiency, this section makes the following improvements to the SA:

- (1) Memorize the best intermediate solution in the search process and update it immediately. The

improvement of memory turns the SA into an intelligent algorithm.

- (2) Set up two thresholds, internal cycle threshold and external cycle threshold, to reduce the computing load while maintaining optimality. The internal cycle threshold refers to the number of cycles that the new solution of continuous disturbance does not generate a better solution at a certain temperature, while the external cycle threshold refers to the number of cycles that the new solution generated by continuous cooling does not generate a better solution. The two thresholds are determined as follows.

First, determine whether the number of internal cycles reaches the threshold; if yes, lower the temperature by one step; otherwise, judge if it conforms to Markov chain. If not, reconduct the process of state generation, state acceptance, and algorithm termination; otherwise, lower the temperature by one step. Second, determine whether the number of external cycles reaches the threshold; if yes, terminate the algorithm and obtain the final solution; otherwise, judge whether the algorithm meets the termination conditions. If yes, terminate the algorithm and obtain the final solution; otherwise, reconduct the process of state generation, state acceptance, and algorithm termination. Terminate the algorithm once the number of iterations  $i$  surpasses the prior fixed constant  $MaxIter$ . In our experiments,  $MaxIter$  was set to  $10^4$ .

The flow of the improved SA is shown in Figure 3.

## 4. Experiments

To test the effectiveness of SA-based algorithm, we first evaluate the control factors of the SA and suggest a good parameter setting. Then, the solution quality and efficiency of the ISA were verified through several experiments. The algorithms are implemented in our previously published carrier-based flight operations simulation [28], which is written by C++ and ran on Microsoft Windows operating system with 4 GB RAM and dual core CPU.

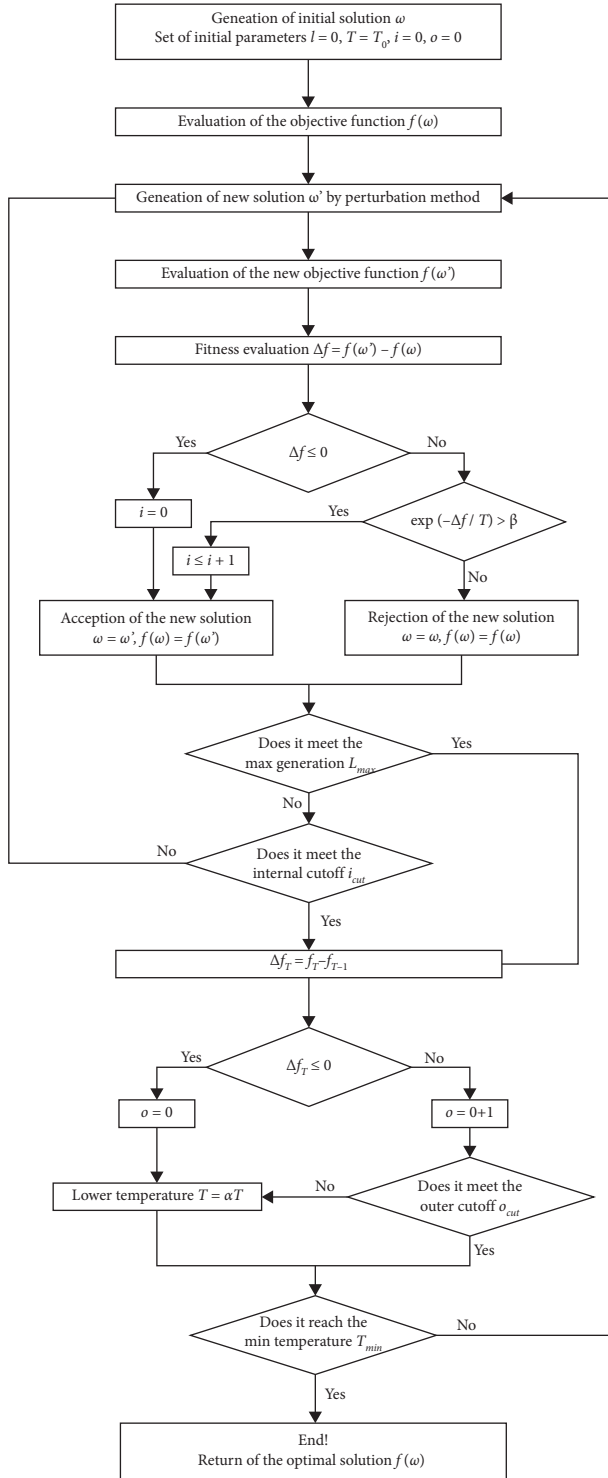


FIGURE 3: Flow of the improved SA.

**4.1. Parameter Tuning.** The efficiency of the SA algorithm is greatly affected by the design of parameters and operators. The full factorial design tests all possible combinations. Such an approach becomes too laborious in the face of numerous factors. Taguchi utilized orthogonal arrays to examine lots of decision variables in a few tests [29] and measured the importance of each factor by its influence on algorithm

performance, using the signal-to-noise ratio:  $10 \log_{10} (\text{objective})^2$ . Following Taguchi's method, the SA control factors were configured as follows: initial solution, initial temperature, cooling rate, and number of neighborhood searches in every temperature. Table 2 shows the different levels of these factors.

Hence, the SA has one 3-level factor and three 4-level factors. The best design among the orthogonal arrays is L16. Thus, additional transform was performed to fit L16 (Table 3).

The relative percentage deviation (RPD) was also adopted to measure the performances:

$$\text{RPD} = \frac{\text{Alg}_{\text{sol}} - \text{Min}_{\text{sol}}}{\text{Min}_{\text{sol}}} \cdot 100\%, \quad (11)$$

where the best solution obtained for one instance is denoted as  $\text{Min}_{\text{sol}}$ , while the objective value is marked as  $\text{Alg}_{\text{sol}}$ . Table 4 lists the S/N ratios and RPD values of each level of the factor value. The results show that A(3), B(4), C(4), and D(4) are the best levels of the factors.

**4.2. Experimental Settings.** The size of the test instances was set to  $n = \{15, 30, 45, 60\}$  tasks, which corresponds to a common mission of strike sorties of 5, 10, 15, and 20 aircraft, respectively. The processing time for jobs on each machine was generated by triangle distribution with mean time according to [30]. There are resource constraints of five types of weapons, four lower-stage elevators, 10 assembling crews, four upper-stage elevators, and six loading crew members, see Tables 5–9 for comprehensive data.

**4.3. Makespan Analysis.** The test of problem uses an ordinance loading plan, ranging from 5 to 20 aircraft. At first, the experiment tries to solve a standard small size problem with optimal solution and preliminarily demonstrates the adaptability and feasibility of the ISA. Then, the proposed ISA was adopted to solve larger size problems and compared with the other methods to reveal its superiority.

To compare the ISA with the SA, control factors were configured as those in the preceding section. The permutation of machines was arranged in ascending order. The initial temperature  $T_0 = 10^3$ , final temperature  $\varepsilon = 0.5$ , cooling rate  $\alpha = 0.99$ , and length of Markov chain  $L = 2000$ .

The base case includes 15 tasks. As the temperature declined (Figure 4(a)), the ISA converged to the optimal solution in 2.705 s, faster than the SA, which converged to the optimal solution in 2.965 s (Figure 5(a)). Although the initial scheduling time of the ISA was higher than that of SA, its faster convergence procedure suggests the good performance of the ISA. Note that the computing time of the ISA was 85.3637 s, much shorter than that (143.7861 s) of the SA.

Next, the ISA performance on larger problems was tested, whereas more tasks bring a greater computing effort for the heuristic. Table 10 lists the average scheduling time and computing time of different tasks, with each task running for 30 times. Figure 6 shows the corresponding plot. The results showed that the ISA achieved a shorter

TABLE 2: The SA control factors.

Factors	Levels	Types
Initial solution (A)	3	(A1) randomness (A2) task-based (A3) ascending order
Initial temperature (B)	4	B1 = 50, B2 = 100, B3 = 500, B4 = 1000
Cooling rate (C)	4	C1 = 0.85, C2 = 0.9, C3 = 0.95, C4 = 0.99
Markov chain (d)	4	D1 = 100, D2 = 500, D3 = 1000, D4 = 2000

TABLE 3: Orthogonal array of L16 of our algorithm test.

Experiment	A	B	C	D
1	1	2	3	3
2	2	4	1	2
3	3	4	3	4
4	1	2	1	1
5	1	3	1	4
6	2	1	3	1
7	3	1	1	3
8	2	3	3	2
9	1	1	4	2
10	2	3	2	3
11	3	3	4	1
12	3	1	2	4
13	1	4	2	1
14	2	2	4	4
15	3	2	2	2
16	1	4	4	3

TABLE 4: Results of the orthogonal test.

Factor with level	Completion time	
	Mean S/N ratio	Mean RPD
Random initial (A1)	-73.6779	2.399
Descending initial (A2)	-73.6634	2.216
Ascending initial* (A3)	-73.6504*	2.060*
T0 = 50 (B1)	-73.6656	2.239
T0 = 100 (B2)	-73.6638	2.223
T0 = 500 (B3)	-73.6669	2.252
T0 = 1000* (B4)	-73.6629*	2.212*
Alpha = 0.85 (C1)	-73.7266	2.962
Alpha = 0.9 (C2)	-73.6791	2.397
Alpha = 0.95 (C3)	-73.6547	2.113
Alpha = 0.99* (C4)	-73.5987*	1.453*
L = 100 (D1)	-73.7602	3.360
L = 500 (D2)	-73.6716	2.310
L = 1000 (D3)	-73.6344	1.869
L = 2000* (D4)	-73.5930*	1.387*

TABLE 5: Ordnance loading plan.

Ordnance type	Spot task									
	A <sub>1</sub>	A <sub>2</sub>	A <sub>3</sub>	A <sub>4</sub>	A <sub>5</sub>	A <sub>6</sub>	A <sub>7</sub>	A <sub>8</sub>	A <sub>9</sub>	A <sub>10</sub>
1	4(2)									
2		4(2)								
3			2(1)	2(1)	2(1)	2(1)	2(1)	2(1)	2(1)	2(1)
4			2(1)	2(1)	2(1)	2(1)	2(1)	2(1)	2(1)	2(1)
5			6(2)	6(2)	6(2)	6(2)				
6							2(2)	2(2)	2(2)	2(2)

TABLE 6: Ordnance assembling time (min).

Ordnance type	1	2	3	4	5	6
Loading time/skid	13	8	7	7	11	6

TABLE 7: Ordnance transport time on flight deck.

Time (s) Upper-stage elevator number	Spot									
	A <sub>1</sub>	A <sub>2</sub>	A <sub>3</sub>	A <sub>4</sub>	A <sub>5</sub>	A <sub>6</sub>	A <sub>7</sub>	A <sub>8</sub>	A <sub>9</sub>	A <sub>10</sub>
1	245	199	148	30	61	107	214	213	402	342
2	106	60	190	92	66	67	96	159	345	286
3	381	337	262	166	127	78	74	91	267	208
4	443	398	320	228	189	134	118	80	217	161

TABLE 8: Ordnance loading time (s).

Ordnance type	1	2	3	4	5	6
Time (skid/s)	660	600	480	480	780	360

TABLE 9: Ordnance crew walking time (s).

Spot	Spot									
	A <sub>1</sub>	A <sub>2</sub>	A <sub>3</sub>	A <sub>4</sub>	A <sub>5</sub>	A <sub>6</sub>	A <sub>7</sub>	A <sub>8</sub>	A <sub>9</sub>	A <sub>10</sub>
A <sub>1</sub>	0	47.53946	168.2	222.6	266.9	331	366	446.9	639	578.8
A <sub>2</sub>	47.5	0	142.9	181.2	225.6	289.4	324.9	404.6	596.4	536.1
A <sub>3</sub>	168.2	142.9	0	97.8	131.5	189.8	220	302.5	492	432.9
A <sub>4</sub>	222.6	181.2	97.8	0	44.4	108.5	143.7	224.3	416.5	356.3
A <sub>5</sub>	266.9	225.6	131.5	44.4	0	64.1	99.3	180.1	372.3	312.1
A <sub>6</sub>	331	289.4	189.9	108.5	64.1	0	36.1	116	308.2	278.2
A <sub>7</sub>	366	324.9	220	143.7	99.3	36.1	0	82.6	274	214.1
A <sub>8</sub>	446.9	404.6	302.5	224.3	180.1	116	82.6	0	192.2	132
A <sub>9</sub>	639	596.4	492	416.5	372.3	308.1623	274.0073	192.1666	0	60.3
A <sub>10</sub>	578.8	536.1	432.9	356.3	312.1	278.2	214.1	132	60.3	0

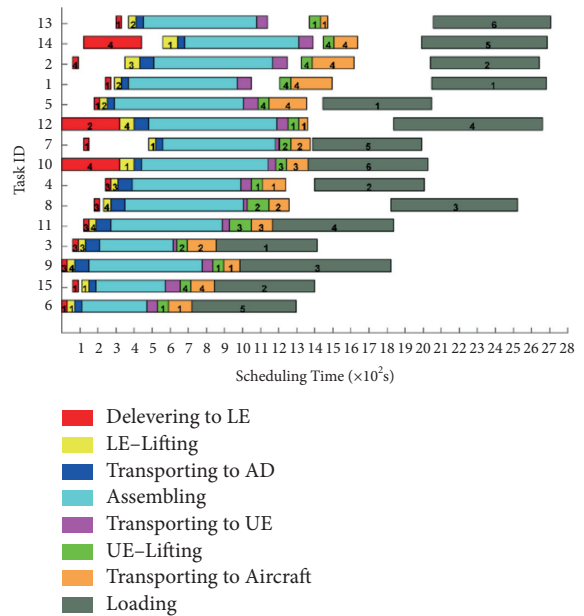
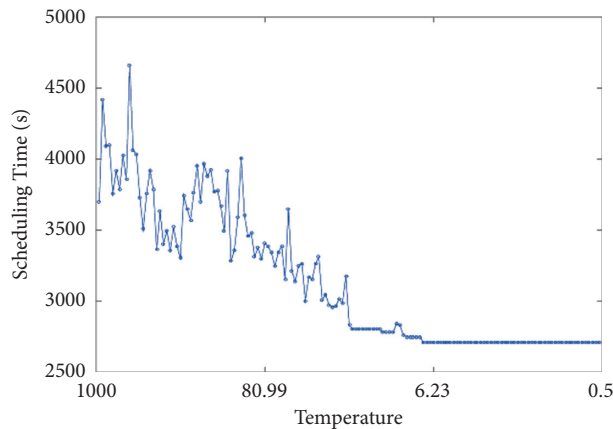


FIGURE 4: Optimal schedule of 15 tasks (5 aircraft) derived by the ISA. (a) Convergence curve. (b) Gantt chart.

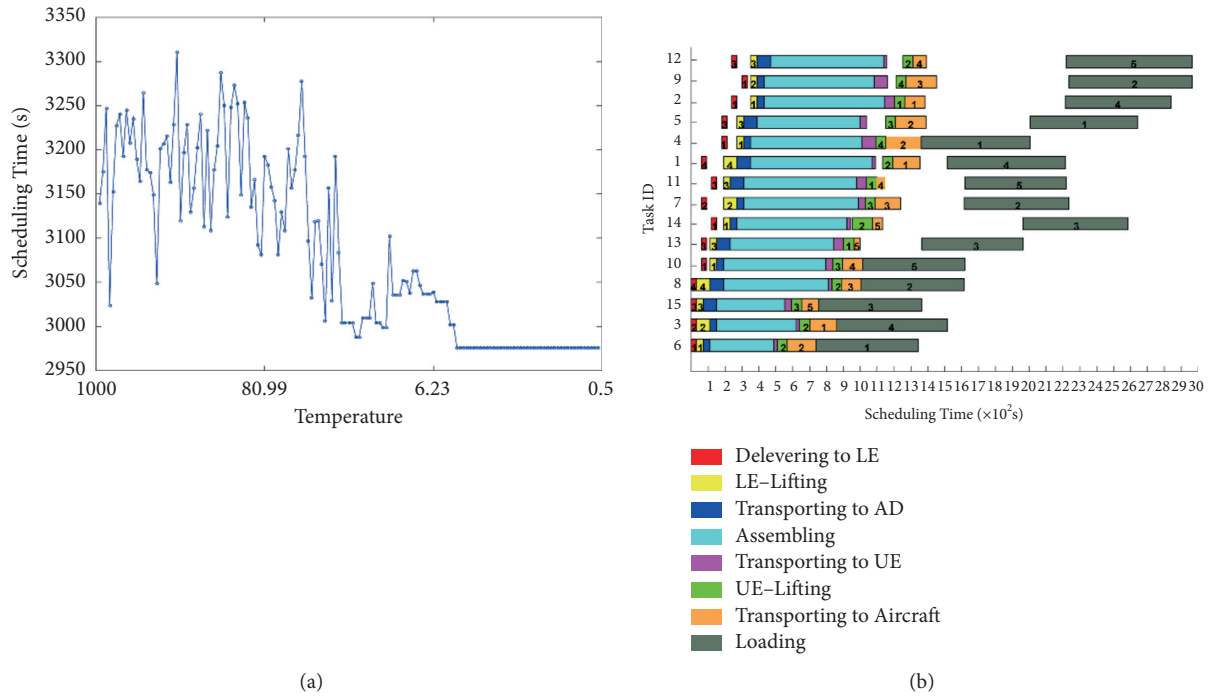


FIGURE 5: Optimal schedule of 15 tasks (5 aircraft) derived by the SA. (a) Convergence curve. (b) Gantt chart.

TABLE 10: Comparison between the ISA and the SA.

Task numbers	Average scheduling time (s)		Average computing time (s)	
	ISA	SA	ISA	SA
15	2965	3100	60	150
30	4200	4500	150	260
45	6100	6500	200	370
60	8000	8600	260	450

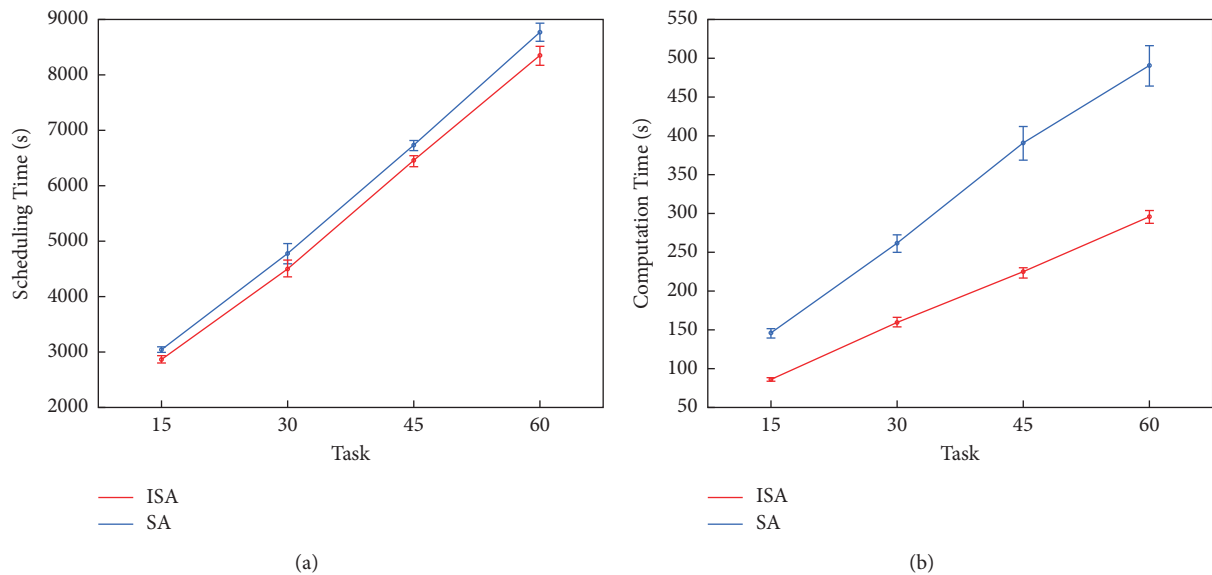


FIGURE 6: Performance of the ISA and the SA on medium and large problems. (a) Scheduling time. (b) Computing time.

scheduling time and a faster computing speed than the SA. Comparing with the SA, the average scheduling time of the ISA reduces from 300 to 600 seconds in each ordnance turnaround cycle for 30 to 60 tasks and the average computing time saves from 110 to 190 seconds. Note that there are usually ten or more cycles in a general flight day; the scheduling time saved by implementing ISA equals one more group sortie generation, which in turn enhanced the firepower capacity of carrier air wing.

The results also show that the bottleneck of ordnance handling is the loading process, where the number of loading crews heavily influences the aircraft turnaround time. When there is high intensity of surge operations, more loading crews should be arranged to handle ordnances.

## 5. Conclusions

This paper treats the aviation ordnance scheduling problem under the HFS framework with multistages, independent parallel machines, and several processing constraints. The simulated annealing algorithm was modified with dual threshold selection to generate faster and better schedules using the proposed matrix perturbation method that keeps the SA independent of the heuristic schemes. The influencing parameters of the improved algorithm are carefully tuned by Taguchi's method. The experimental results demonstrate the effectiveness of ISA, which provides a practical solution to a broad application in dealing with stochastic hybrid flow shop scheduling problems.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References

- [1] A. Jewell, *Sortie Generation Capacity of Embarked Airwings*, Center For Naval Analyses Alexandria Va, Arlington, VA, USA, 1998.
- [2] Naval Air Systems Command, *CV Naval Air Training and Operating Procedures Standardization*, NATOPS Manual, Patuxent River, MD, USA, 2009.
- [3] R. Ruiz and J. A. Vázquez-Rodríguez, "The hybrid flow shop scheduling problem," *European Journal of Operational Research*, vol. 205, no. 1, pp. 1–18, 2010.
- [4] H. Emmons and G. Vairaktarakis, *Flow Shop Scheduling: Theoretical Results, Algorithms, and Applications*, Springer Science & Business Media, Berlin, Germany, 2012.
- [5] E. Mokotoff, "Parallel machine scheduling problems: a survey," *Asia-Pacific Journal of Operational Research*, vol. 18, no. 2, p. 193, 2001.
- [6] M. M. Ahmadian, M. Khatami, A. Salehipour, and T. C. E. Cheng, "Four decades of research on the open-shop scheduling problem to minimize the makespan," *European Journal of Operational Research*, vol. 259, no. 2, pp. 399–426, 2021.
- [7] A. O. Bedhief and N. Dridi, "A genetic algorithm for three-stage hybrid flow shop scheduling problem with dedicated machines," *Journal Européen des Systèmes Automatisés*, vol. 53, no. 3, pp. 357–368, 2020.
- [8] J. N. D. Gupta, "Two-stage, hybrid flowshop scheduling problem," *Journal of the Operational Research Society*, vol. 39, no. 4, pp. 359–364, 1988.
- [9] G.-C. Lee and Y.-D. Kim, "A branch-and-bound algorithm for a two-stage hybrid flowshop scheduling problem minimizing total tardiness," *International Journal of Production Research*, vol. 42, no. 22, pp. 4731–4743, 2004.
- [10] T. Sawik, "Integer programming approach to reactive scheduling in make-to-order manufacturing," *Mathematical and Computer Modelling*, vol. 46, no. 11–12, pp. 1373–1387, 2007.
- [11] I. Ribas, R. Leisten, and J. M. Framiñan, "Review and classification of hybrid flow shop scheduling problems from a production system and a solutions procedure perspective," *Computers & Operations Research*, vol. 37, no. 8, pp. 1439–1454, 2010.
- [12] G.-C. Lee, "Estimating order lead times in hybrid flowshops with different scheduling rules," *Computers & Industrial Engineering*, vol. 56, no. 4, pp. 1668–1674, 2009.
- [13] G. Bektur and T. Saraç, "A mathematical model and heuristic algorithms for an unrelated parallel machine scheduling problem with sequence-dependent setup times, machine eligibility restrictions and a common server," *Computers & Operations Research*, vol. 103, pp. 46–63, 2019.
- [14] T. Yang, Y. Kuo, and C. Cho, "A genetic algorithms simulation approach for the multi-attribute combinatorial dispatching decision problem," *European Journal of Operational Research*, vol. 176, no. 3, pp. 1859–1873, 2007.
- [15] D. Li, C. Liu, and K. Li, "A remanufacturing logistics network model based on improved multi-objective ant colony optimization," *Journal Européen des Systèmes Automatisés*, vol. 52, no. 4, pp. 391–395, 2019.
- [16] F. Glover and M. Laguna, "Tabu search, handbook of combinatorial optimization," in *Handbook of Combinatorial Optimization*, pp. 2093–2229, Springer, Boston, MA, USA, 1998.
- [17] C. B. Rabah, G. Coatrieux, and R. Abdelfattah, "Boosting up source scanner identification using wavelets and convolutional neural networks," *Traitement du Signal*, vol. 37, no. 6, pp. 881–888, 2020.
- [18] D. A. B. Fernandes, M. M. Freire, P. A. P. Fazendeiro, and P. R. M. Inácio, "Applications of artificial immune systems to computer security: A survey," *Journal of Information Security and Applications*, vol. 35, pp. 138–159, 2017.
- [19] D. Delahaye, S. Chaimatanan, and M. Mongeau, "Simulated annealing: From basics to applications," *Handbook of Metaheuristics*, vol. 272, pp. 1–35, 2019.
- [20] N. Maaroju, "Choosing the Best Heuristic for a NP-Problem," Doctoral Dissertation, Thapar University, Patiala, Punjab, 2009.
- [21] A. Franzin and T. Stützle, "Revisiting simulated annealing: A component-based analysis," *Computers & Operations Research*, vol. 104, pp. 191–206, 2019.
- [22] K. Walczak and J. Mańdziuk, "Applying hybrid Monte Carlo tree search methods to risk-aware project scheduling problem," *Information Sciences*, vol. 460, pp. 450–468, 2018.
- [23] Commander Naval Air Systems Command, *CV Flight/Hangar Deck NATOPS Manual*, NAVAIR 00-80T-120, Patuxent River, MA, USA, 2005.



- [24] P. C. Goshopn, "Aviation ordnanceman," *Navel Education and Training Professional Development and Technology Center. Navedtra*, vol. 14313, 2001.
- [25] D. Leitold, A. Vathy-Fogarassy, and J. Abonyi, "Empirical working time distribution-based line balancing with integrated simulated annealing and dynamic programming," *Central European Journal of Operations Research*, vol. 27, no. 2, pp. 455–473, 2019.
- [26] S. Rajendran, C. Rajendran, and R. Leisten, "Heuristic rules for tie-breaking in the implementation of the NEH heuristic for permutation flow-shop scheduling," *International Journal of Operational Research*, vol. 28, no. 1, pp. 87–97, 2017.
- [27] B. Maenhout and M. Vanhoucke, "A perturbation mathematical heuristic for the integrated personnel shift and task rescheduling problem," *European Journal of Operational Research*, vol. 269, no. 3, pp. 806–823, 2018.
- [28] N. Wang, X. Meng, Q. Liu, and J. Li, "High level architecture based simulation for aircraft carrier deck operations," in *Proceedings of the 2016 IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*, pp. 765–769, Xi'an, China, October 2016.
- [29] J. D. Kechagias, K.-E. Aslani, N. A. Fountas, N. M. Vaxevanidis, and D. E. Manolakos, "A comparative investigation of Taguchi and full factorial design for machinability prediction in turning of a titanium alloy," *Measurement*, vol. 151, p. 107213, 2020.
- [30] A. Jewell, M. A. Wigge, C. M. Gagnon, L. A. Lynn, and K. M. Kirk, *USS Nimitz and Carrier Airwing Nine Surge Demonstration*, Center for Naval Analyses Alexandria Va, Arlington, VA, USA, 1998.

## Retraction

# Retracted: Research on Probability Distribution of Short-Term Photovoltaic Output Forecast Error Based on Numerical Characteristic Clustering

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] P. Yan, C. Xiang, T. Li et al., "Research on Probability Distribution of Short-Term Photovoltaic Output Forecast Error Based on Numerical Characteristic Clustering," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 5355286, 11 pages, 2022.

## Research Article

# Research on Probability Distribution of Short-Term Photovoltaic Output Forecast Error Based on Numerical Characteristic Clustering

Peng Yan , Chenmeng Xiang , Tiecheng Li , Xuekai Hu , Wen Zhou , Lei Wang ,  
and Liang Meng 

State Grid Hebei Electric Power Research Institute, Shijiazhuang, China

Correspondence should be addressed to Peng Yan; [dyy\\_yanp@163.com](mailto:dyy_yanp@163.com)

Received 22 December 2021; Accepted 15 January 2022; Published 1 February 2022

Academic Editor: Daqing Gong

Copyright © 2022 Peng Yan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The forecast error characteristic analysis of short-term photovoltaic power generation can provide a reliable reference for power system optimal dispatching. In this paper, the total in-day error level was stratified by fuzzy C-means algorithm. Then the historical PV output data based on the numerical characteristics of point prediction output were classified. A General Gauss Mixed Model was proposed to fit the forecast error distribution of various photovoltaic output forecast error distribution. The impact of meteorological factors together with numerical characteristics on the forecast error was taken into full consideration in this analysis method. The predicted point output with high volatility can be accurately captured, and the reliable confidence interval is given. The proposed method is independent of the point prediction algorithm and has strong applicability. The General Gauss Mixed Model can meet the peak diversity, bias, and multimodal properties of the error distribution, and the fitting effect is superior to the normal distribution, the Laplace distribution, and the  $t$  Location-Scale distribution model. The error model has a flexible shape, a concise expression, and high practical value for engineering.

## 1. Introduction

Facing the double pressure of energy crisis and environmental pollution, people pay more and more attention to the new energy generation technology with clean and environmental protection characteristics. Compared with wind power, photovoltaic power generation requires less geographical environment and is more suitable for multiregional promotion and application. However, PV power generation is highly random and intermittent, and large-scale grid connection affects the stability and economy of the system [1]. The accuracy of photovoltaic power prediction has a direct impact on its consumption. Domestic and foreign scholars have conducted relevant studies, and the existing prediction models are divided into two categories: first, direct prediction algorithms such as regression models [2–4], gray prediction models [5–7], neural network models [8–11], and probabilistic models [12] are used; second,

indirect prediction algorithms such as electronic component models [13], simple physical models [14, 15], and complex physical methods [16, 17] are used. The use of different prediction algorithms can have different degrees of prediction errors.

There are only a few literatures on the forecast error of PV power generation at home and abroad, and the description of the prediction error of PV output in some literature is based on the assumption that it obeys normal distribution. The PV output uncertainty needs to be considered when studying the optimal scheduling of power systems, and most of the literature uses the actual output value in the form of the sum of the predicted output and the forecast error. Literature [18] shows that a 10% forecast error produces deviated power exceeding 15% of the rated power value, while a 15% forecast error produces deviated power exceeding 25% of the rated power value, and the forecast error directly affects the safe and stable operation of the

system. Based on the assumption that the forecast error obeys normal distribution, the results obtained in [19–21] are different from the actual statistical results. The research in [22] shows that weather factors have great influence on the forecast error, and the forecast error of solar volts in sunny days is close to normal distribution. The feasibility of using  $t$  Location-Scale model to describe the forecast error of PV output is proposed and verified in [23]. The statistical results show that the PV output forecast error distribution has multiple peaks, while the existing research using single distribution model is weak in describing the multi-peaks. Therefore, [24–26] propose to model the forecast error by Gaussian mixture model (GMM), but the value range of GMM is from negative infinity to positive infinity, which is obviously not applicable for the description of the actual PV output forecast error directly. Literature [27] trains artificial neural networks with a large number of samples to build a forecast error model for photovoltaic power generation, which can avoid the deviation of prediction accuracy caused by model setting and parameter estimation. Literature [28] introduces regularized penalty function and error function to construct the objective function of PV prediction model; the Pearson correlation coefficient between PV power generation and each feature is analyzed, and the abnormal data of the features are also preprocessed. The above studies all focus on the optimization of the model. Because of the random characteristics of meteorological factors such as solar irradiation, temperature, and wind speed, the forecast error of photovoltaic output does not have a certain distribution characteristic, and it is difficult for the established forecast error model to achieve ideal accuracy. The distribution characteristics of PV output forecast error under different meteorological conditions and numerical characteristics cannot be ignored, so it is necessary to cluster the forecast error according to the conditions. At present, there are few researches in this field, so a flexible distribution model is needed, which can meet the requirements of skewness and peak diversity of PV output forecast error.

In this paper, the effects of meteorological and numerical characteristics on the real-time power forecast error of photovoltaic power generation are studied. Based on the corresponding meteorological data, the historical error samples are clustered into three categories by fuzzy C-means clustering, and the error areas are divided into two categories according to the error size. In order to describe the forecast error distribution more accurately, a general Gaussian mixture model based on the traditional Gaussian distribution is proposed. Compared with the traditional Gaussian model, this model can describe the error distribution of different kurtosis and shape more accurately.

In addition, this method is universal and is not affected by photovoltaic power prediction algorithm and the geographical location of photovoltaic power stations.

## 2. Cluster Analysis of Photovoltaic Output Forecast Error

Short-term forecast error of photovoltaic output is mainly affected by weather and numerical characteristics of

prediction points. Among the factors representing weather, weather type, temperature, temperature difference, and wind speed are selected as indicators to analyze the correlation with photovoltaic forecast error. Therefore, firstly, the PV intraday forecast error samples are clustered into three categories according to the weather characteristics, and then the error samples obtained by classification are used as training samples to discriminate the subsequent errors. After determining the classification, the forecast error is divided into large error and small error according to its numerical characteristics. Finally, Gaussian mixture distribution is used for statistical fitting within the class, and a reliable confidence interval is provided for predicting the PV error distribution according to the fitting information.

To determine the confidence interval of photovoltaic error distribution, the steps are shown in Figure 1:

- (1) According to meteorological factors, the historical data of photovoltaic power generation forecast error are clustered into three categories
- (2) Taking amplitude and step size as indexes, the error data in cluster are divided into large error and small error
- (3) The error database will be established according to the error samples clustered by meteorological factors, which is convenient to provide the error interval meeting the error requirements

## 3. Influencing Factors of Photovoltaic Power Forecast Error

Photovoltaic panels absorb solar energy and generate electricity based on Volta effect. Its power generation is affected by meteorological factors, especially illumination and temperature [29]. Literature [30] proposes a photovoltaic power prediction method based on clear coefficient and multilevel similarity matching. In addition, the statistical results show that the forecast error of photovoltaic power generation is directly related to the amplitude and climbing of predicted output. Therefore, this paper studies the factors that affect the error distribution of PV power prediction from two angles of meteorological and numerical factors, which provides important reference information for error discrimination clustering and obtaining reliable confidence intervals.

*3.1. Analysis of the Influence of Meteorological Factors on Forecast Error.* To study the influence of meteorology on forecast error, we should first index meteorological factors concretely. In order to accurately scale meteorological factors, four factors are selected to express: weather type, intraday difference between maximum and minimum temperature, maximum temperature, and wind speed. After that, the influence of these four factors on forecast error is studied, which also provides variables for later error discriminant analysis.

The British statistician R. A. Fisher put forward the variance analysis method in the 1920s. [31]. The variance analysis method can determine the factors that have the

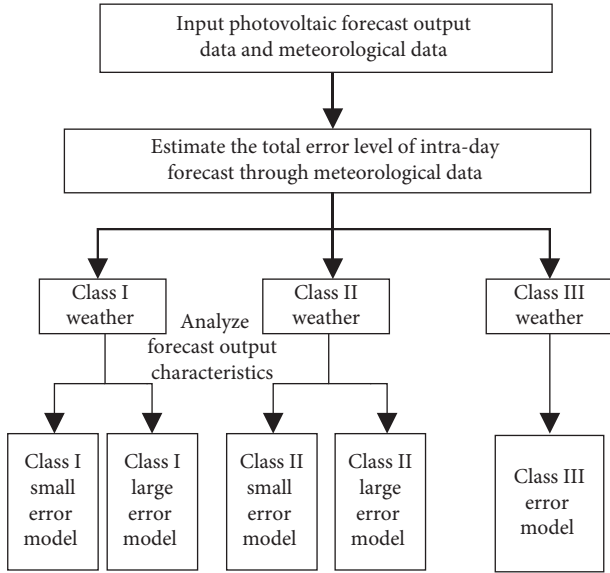


FIGURE 1: Schematic diagram of the research method of photovoltaic output forecast error distribution.

main effect on the target object from many factors. It determines the influence of research elements on the target object by analyzing the contribution of different elements to the overall target. The specific operation process is to analyze the differences between different groups and within groups. The specific discrimination process is as follows:

$$\begin{cases} MSb = \frac{SSb}{dfb}, \\ MSw = \frac{dfb}{dfw}, \end{cases} \quad (1)$$

where  $SSb$  represents the intergroup differences;  $SSw$  represents intragroup differences;  $dfb$  and  $dfw$  are the degrees of freedom between groups and within groups, respectively. Whether the experimental factors have obvious influence on the research object is judged by the ratio of  $MSb/MSw$  and the  $F$  distribution composed of  $MSb/MSw$ . The probability  $P$  value of  $F$  value greater than a specific value under the test hypothesis can be obtained by consulting the  $F$  boundary value table. Select 0.05 as the test critical value. When  $P < 0.05$ , it is considered that the test factors have significant differences on the research objects; otherwise, it is considered that there is no obvious influence. When studying the influence of weather factors on the forecast error of photovoltaic power generation, the selected test factors and levels are shown in Table 1.

The influence of meteorological factors on PV forecast error is analyzed. Firstly, the meteorological factors are indexed as weather type  $A$ , intraday temperature difference  $B$ , intraday maximum temperature  $C$ , and wind speed grade  $D$ . Photovoltaic forecast error is quantified by sum of squares of errors (DSSE), and weather types are quantified by sunny degree assignment [1–3]. Taking PV in Brussels area in 2016 as an example, the results of the analysis of variance are shown in Table 2.

In Table 2, the main effect of four variables and the interaction effect between two variables are selected as factors, and the sum of squares of variance, degree of freedom, mean square, observed value of  $F$  distribution, and test  $P$  value are used as indexes for analysis. As can be seen from Table 2, the  $P$  values of principal factor  $B$ , principal factor  $C$ , and interactive factor  $B^*C$  are less than 0.05. That is to say, at the significant level of 0.05, the effects of principal factor  $B$ , principal factor  $C$ , and interactive factor  $B^*C$  are significant. At the significant level of 0.05, other factors are not significant. From the results, we can see that, among the single factors selected in the early stage, factor  $D$  has the least significant influence on the error. In order to remove its influence on other factors and extract the components more accurately, factor  $D$  is removed and then does variance analysis again. The results are shown in Table 3.

As can be seen from Table 3, after removing the influence of factor  $D$ , the influence of factors  $A$ ,  $B$ , and  $C$  is more significant. At a significant level of 0.05, weather type, intraday temperature difference, maximum temperature, and the interaction between intraday temperature difference and maximum temperature have the most significant influence on the total forecast error level.

**3.2. Analysis of the Influence of Numerical Characteristics of Photovoltaic Output on Forecast Error.** Photovoltaic panels usually run in the maximum power tracking state. When external factors such as illumination and temperature change, the controller controls the operating point of PV array to change, so the forecast error of photovoltaic output is related to the performance of the controller. The prediction power amplitude is selected as factor  $E$ , and the adjacent prediction output difference is factor  $G$ , and the influence of the two factors on the short-term photovoltaic output forecast error is analyzed. The rated capacity of two factors is taken as the reference value to make the standard output, and the specific level values are shown in Table 4.

Based on the photovoltaic power generation data of Brussels region in Belgium in 2016, the output amplitude and climbing power are used as indexes for principal component analysis. The results are shown in Table 5.

At a significant level of 0.05, all factors in Table 5 passed the test. Therefore, it can be seen that both the amplitude of photovoltaic output and climbing power have a significant impact on the forecast error.

**3.3. Cluster Analysis of Influencing Factors of Photovoltaic Forecast Error.** From the above analysis, it can be seen that there are many factors affecting photovoltaic forecast error. In order to facilitate the subsequent study of forecast error, it is necessary to reduce the variable dimension. In this paper, the fuzzy C-means clustering method is used to cluster the historical data DSSE, and the meteorological data are classified according to the clustering results, which can be used to discriminate and analyze the meteorological types of the forecast days and estimate the total forecast error level of the day.

TABLE 1: List of factors and levels.

Level	Factor			
	A (weather)	B (temperature difference (°C))	C ( $T_{\max}$ (°C))	D (wind speed grade)
1	Sunny day	1~5	-1~10	1~2
2	Cloudy	6~10	11~20	3~4
3	Rainy day	11~15	21~32	5~6

TABLE 2: Factors and test parameter values.

Source	Sum sq.	d.f.	Mean sq.	$F$	$P$
A	0.324	2	0.1594	1.29	0.2762
B	1.301	2	0.6488	5.26	0.0051
C	2.529	2	1.2646	10.27	$\leq 0.0001$
D	0.100	2	0.0501	0.41	0.6663
A*B	0.270	4	0.0676	0.55	0.7011
A*C	0.554	4	0.1384	1.12	0.3453
B*C	2.909	4	0.7273	5.9	0.0001
Error	41.141	334	0.1232		
Total	51.560	354			

TABLE 3: Processed factors and processed test parameter values.

Source	Sum sq.	d.f.	Mean sq.	$F$	$P$
A	1.133	2	0.5650	4.62	0.0117
B	1.612	2	0.7983	6.57	0.0019
C	2.607	2	1.3001	10.77	$\leq 0.0001$
A*C	0.655	4	0.1664	1.36	0.2886
B*C	3.027	4	0.7374	6.16	0.0001
Error	43.225	340	0.1225		
Total	52.137	354			

Fuzzy C-mean clustering method is used in cases where there are no clear boundaries between the classified objects. Therefore, fuzzy C-means clustering method is used to combine the meteorological factors obtained above into three categories, namely, Class I, Class II, and Class III. Taking the total error level of photovoltaic prediction DSSE as the error index, the observation matrix is listed in days:

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{np} \end{pmatrix}, \quad (2)$$

where each row of  $X$  is a sample of one day and each column has  $p$  observations within one day; i.e.,  $X$  is a matrix consisting of observations of  $p$  variables over  $n$  days;  $X_{np}$  represents the observed value of the  $p$ -th variable on the  $n$ -th day;  $n$  samples are divided into  $c$  classes ( $2 \leq c \leq n$ ) and  $V = \{v_1, v_2, \dots, v_c\}$  is recorded as  $c$  cluster centers. Samples  $x_k$  are not strictly divided into a certain class but belong to a certain class by membership degree  $u_{ik}$ , and  $0 \leq u_k \leq 1, \sum_{i=1}^c u_{ik} = 1$ . Define the target function:

$$J(U, V) = \sum_{k=1}^n \sum_{i=1}^c u_{ik}^m d_{ik}^2, \quad (3)$$

where  $U = (u_{ik})_{c \times n}$  is the membership matrix;  $d_{ik} = \|x_k - v_i\|$ .  $J(U, V)$  represents the sum of weighted

square distances from samples to cluster centers in each class. Based on fuzzy C-means clustering method, Lagrange multiplier method [32] and iterative method [31] are often used to solve the objective function to obtain the minimum values of  $U$  and  $V$ .

Fuzzy C-means clustering method is used to cluster photovoltaic short-term forecast errors. The results are shown in Figure 2, where dots represent error samples. It can be seen from the figure that all error samples are clustered into three classes, and Class I error is the smallest, Class III error is the largest, and Class II error is moderate. After getting the error clustering results, the corresponding meteorological data are also classified and archived and used as their own training samples to discriminate and analyze the weather on the forecast day.

Figure 3 shows the percentage of sunny, rainy, and snowy weather on the left side and the sample mean values of intraday temperature difference, maximum temperature, and minimum temperature on the right side, which shows the clustering of meteorological data according to DSSE value clustering date. As can be seen from the above figure, the proportion of various weather types of Class I weather and Class II weather is similar, but the temperature of Class I weather is low and the temperature difference is small. The intraday temperature and temperature difference of Class II weather and Class III weather are similar, but cloudy days account for a high proportion and sunny and rainy days account for a small proportion in Class III weather.



TABLE 4: List of factors and levels.

Level	Factor	
	<i>E</i> (day-ahead forecast output)	<i>G</i> (step length)
1	0~0.105	-0.06~-0.045
2	0.105~0.21	-0.045~-0.03
3	0.21~0.315	-0.03~-0.015
4	0.315~0.42	-0.015~-0.003
5	0.42~0.525	-0.003~0.009
6	0.525~0.63	0.009~0.021
7	0.63~0.735	0.021~0.033
8	0.735~0.84	0.033~0.045

TABLE 5: Factors and test parameter values.

Source	Sum sq.	d.f.	Mean sq.	<i>F</i>	<i>P</i>
<i>E</i>	0.309	5	0.0539	9.36	≤0.0001
<i>G</i>	0.312	3	0.1801	15.69	≤0.0001
<i>E</i> * <i>G</i>	1.567	43	0.0361	6.01	≤0.0001
Error	306.495	50982	0.0060		
Total	318.504	51039			

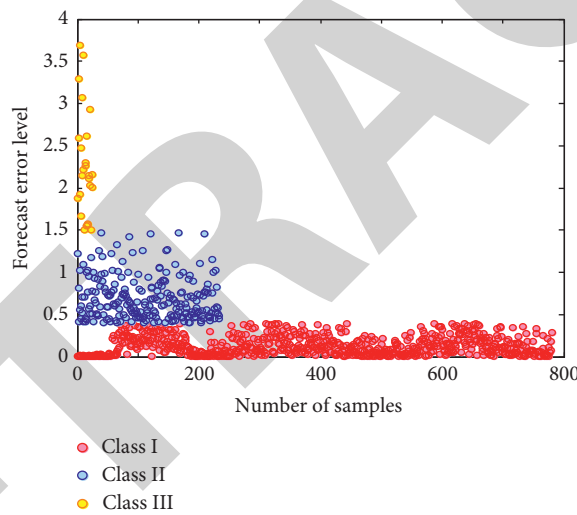


FIGURE 2: Fuzzy C-means clustering results.

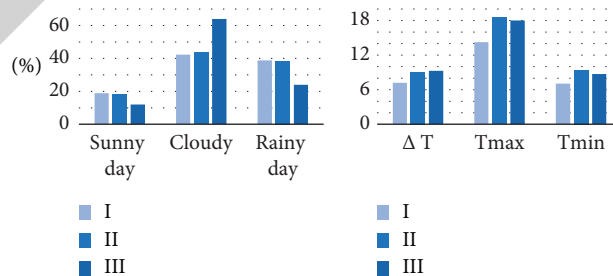


FIGURE 3: Clustering results of meteorological data.

In order to get the weather category of the forecast day, it is necessary to train each group of meteorological data as samples. In the training process, the intraday temperature difference range is [0°C, 18°C], and the intraday maximum

temperature range is [-3°C, 34°C]. Mahalanobis distance, proposed by Indian statistician P.C. Mahalanobis, is a measure of similarity between two points in multidimensional space, which can effectively calculate the similarity

between two unknown sample sets. Different from Euclidean distance, Mahalanobis distance between two points is independent of the measurement unit of the original data and is not affected by dimension. It can be seen from formula (4) that Mahalanobis distance is the product of Euclidean distance and spatial covariance inverse matrix. When the covariance matrix is unit matrix, Mahalanobis distance degenerates to Euclidean distance. For the factors with obvious differences, Mahalanobis distance is used to calculate the similarity, as shown in the following formula:

$$d^2(x, y) = (x - y)^T \Sigma^{-1} (x - y). \quad (4)$$

**3.4. Classification Processing of Forecast Error.** The research results in Section 3.2 of this paper show that the amplitude and step size of the predicted output have a significant interaction. In Section 3.2, the mean absolute error (MAE) of the samples combined by two factors at different levels is counted. The results are shown in Figure 4, and the data values are detailed in Table 6.

The statistical situation in Figure 4 is classified and described as three cases: in case 1, the combination of  $E$  and  $G$  values is missing in the lower left corner and the lower right corner of the figure, that is,  $\{7, 1\}$ ,  $\{8, 1\}$ ,  $\{8, 2\}$ ,  $\{8, 7\}$ ,  $\{7, 8\}$ ,  $\{8, 8\}$  combined samples; in case 2, the dotted box area with the highest heat in the middle of Figure 2 is a large error area  $E \in [3, 6]$  and  $G \in [3, 6]$ ; in case 3, the area that belongs neither to the large error area nor to the missing area is annularly distributed around the large error area, which is defined as the small error area.

Based on the clustering results of meteorological data, according to the characteristics of prediction output amplitude and step size, the historical data of Class I and Class II forecast errors are further divided into small error area and large error area; Class III error itself has high uncertainty and less samples, so it is no longer classified.

## 4. Forecast Error Model of Short-Term Photovoltaic Power Generation Output

**4.1. General Gaussian Mixture Model.** The statistical distribution of PV short-term output forecast error has the characteristics of asymmetry, diverse kurtosis, and multiple peaks. The traditional probability density function of Gaussian mixture distribution is defined as formula (5), where the sum of coefficients of each Gaussian term is 1.

$$f(x|\theta) = \sum_{k=1}^n a_k \phi(x|\theta_k), \quad (5)$$

where  $a_k$  is the weighting factor,  $a_k \geq 0$ ,  $\sum_{k=1}^n a_k = 1$ ;  $\theta_k = (\mu_k, \sigma_k^2)$ ;  $\phi(x|\theta_k)$  is Gaussian distribution function as shown in the following formula:

$$\phi(x|\theta_k) = \frac{1}{\sqrt{2\pi}\sigma_k} \exp\left(-\frac{(x - \mu_k)^2}{2\sigma_k^2}\right), \quad (6)$$

and its cumulative distribution function is

$$F(x|\theta) = \sum_{k=1}^n a_k \int_{-\infty}^x \phi(x|\theta_k) dt. \quad (7)$$

The random variable range of Gaussian mixture distribution is  $(-\infty, +\infty)$ , but the short-term forecast error of photovoltaic is not the same in practice. To solve this problem, a general Gaussian mixture model (GGMM) is proposed based on the traditional Gaussian mixture distribution. The definition formula of GGMM is basically the same as the traditional Gaussian mixture distribution, except that there is no strict and unique restriction on the sum of the weight coefficients of each Gaussian term. Theoretically, the proposed general Gaussian mixture model is more flexible than the traditional Gaussian mixture model, and it is more applicable to describe the short-term photovoltaic output with asymmetric and multiplex characteristics.

**4.2. Model Parameter Estimation and Accuracy Evaluation.** In this paper, the least square method is used as the main method to estimate the model parameters, and the estimated parameters are obtained by the nonlinear curve fitting function `lsqcurvefit` in MATLAB. Multivariate determination coefficient ( $R^2$ ) is also called goodness of fit, and its value determines the close degree of correlation. When  $R^2$  is closer to 1, the reference value of related equations is higher. On the contrary, the closer it is to 0, the lower the reference value. Root mean square error (RMSE), also called standard error, is very sensitive to a set of extra-large or extra-small errors in fitting, so it can well reflect the precision of fitting. The closer RMSE is to 0, the higher the fitting precision is. The calculation formula is as follows:

$$\begin{cases} R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}, \\ RMSE = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{i}} \end{cases} \quad (8)$$

where  $y_i$  is the actual statistical probability density,  $\hat{y}_i$  is the curve fitting value,  $\bar{y}$  is the average value, and subscript  $i$  represents the  $i$ - the error interval.

## 5. Example Analysis

In order to verify the effectiveness and applicability of the proposed method, the historical data of PV short-term prediction in Brussels, Belgium, is used as an example to simulate in MATLAB software. Among them, the historical data from 2014 to 2016 are used as training samples to establish the forecast error model, and some data from 2017 are selected as test data to test the accuracy of the model. The data in this article comes from the official website of Elia, Belgium.

Elia official website makes the next day's output forecast at 11:00 a.m. every day and updates the next day's 24-hour (96 o'clock) output at 11:45 a.m., with a time resolution of point/15 min. The collected photovoltaic output data and

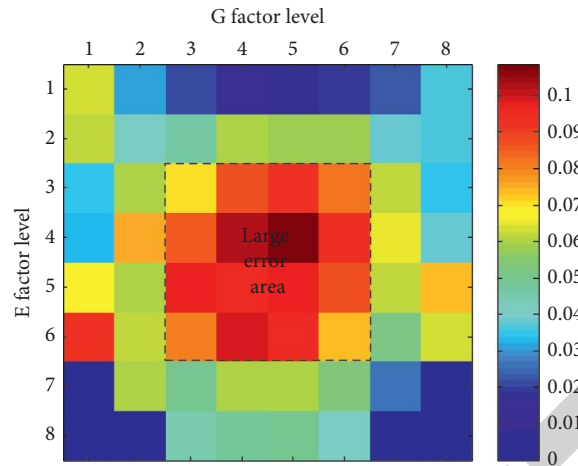


FIGURE 4: MAE statistical chart of samples at each level of  $E$  and  $G$  factors.

TABLE 6: The MAE statistics table of the samples of  $E$  and  $G$  factors at each level.

$E$	$G$							
	1	2	3	4	5	6	7	8
1	0.061	0.019	0.013	0.013	0.013	0.013	0.013	0.032
2	0.061	0.032	0.048	0.061	0.061	0.061	0.032	0.032
3	0.032	0.061	0.072	0.092	0.092	0.084	0.061	0.032
4	0.019	0.084	0.092	0.101	0.115	0.092	0.072	0.032
6	0.072	0.061	0.101	0.101	0.101	0.092	0.061	0.072
6	8	0.061	0.084	0.101	0.101	0.072	0.048	0.061
7	0	0.061	0.048	0.061	0.061	0.048	0.019	0
8	0	0	0.048	0.048	0.048	0.032	0	0

TABLE 7: Accuracy evaluation of different models.

	Model	$R^2$	RMSE
Class I small error	3GGMM	0.9856	0.0756
	Laplace	0.9326	0.8993
	$t$ -distribution	0.9795	0.6865
	Normal distribution	0.8003	1.8454
Class II small error	3GGMM	0.9991	0.1027
	Laplace	0.5173	1.9803
	$t$ -distribution	0.9675	0.4952
	Normal distribution	0.7264	1.5314
Class III small error	3GGMM	0.9023	0.3785
	Laplace	0.4802	0.8695
	$t$ -distribution	0.7203	0.6263
	Normal distribution	0.3254	0.9886
Class I large error	3GGMM	0.9995	0.1132
	Laplace	0.9348	0.4165
	$t$ -distribution	0.9951	0.1403
	Normal distribution	0.9951	0.1406
Class II large error	3GGMM	0.9601	0.1385
	Laplace	0.7784	0.4625
	$t$ -distribution	0.8728	0.3098
	Normal distribution	0.8756	0.3178

meteorological data have the problems of missing data and abnormal data. For the lack of intraday meteorological data, the output data of the solar photovoltaic system will not be

used. And when either the predicted data or the measured data is missing and cannot be repaired, the data will not be used.

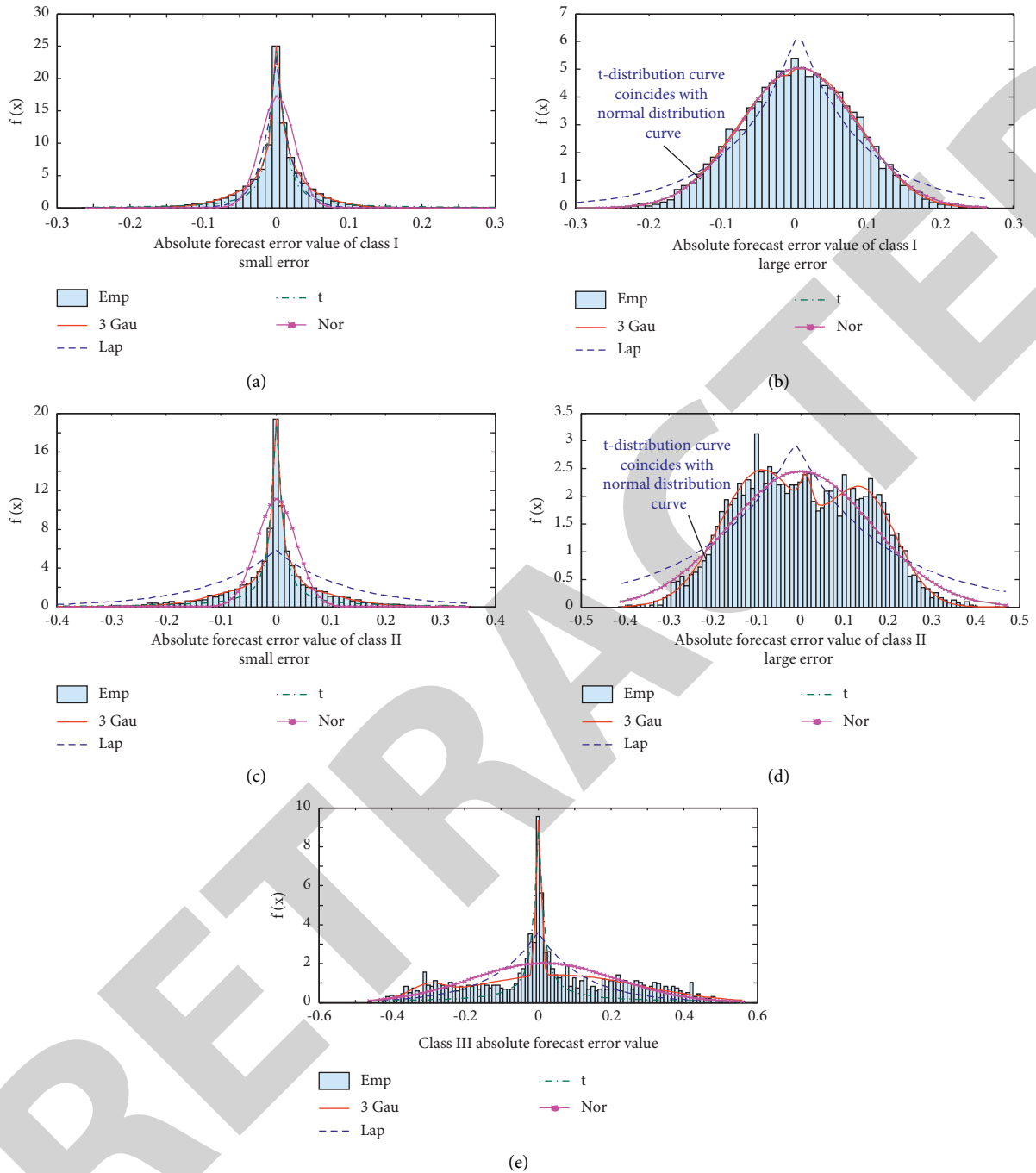


FIGURE 5: Comparison chart of distribution fitting of five groups of errors.

**5.1. Comparison of Model Accuracy.** In order to verify the accuracy and superiority of the model, the PV forecast error distribution model commonly used in the existing literature is used for comparison. The detailed fitting results of each model are shown in Table 7, and the fitting results are shown in Figure 5. In the figure, Emp represents the original error statistical results, 3Gau represents the proposed third-order general Gaussian mixture distribution, Lap represents Laplace distribution,  $t$  represents  $t$  Location-Scale distribution, and Nor represents normal distribution.

It can be seen from the results in Figure 5 that when the fitting distribution presents Class I and Class II small errors with higher peak degree, the accuracy of normal distribution is the lowest, followed by Laplace and Location-Scale distribution, and the proposed general Gaussian mixture distribution has the best effect. Normal distribution is obviously not enough to track spikes. When the fitting distribution shows large errors of Class I and Class II with gentle kurtosis, the effects of the three distributions mentioned above are not comparable to those of the general Gaussian mixture

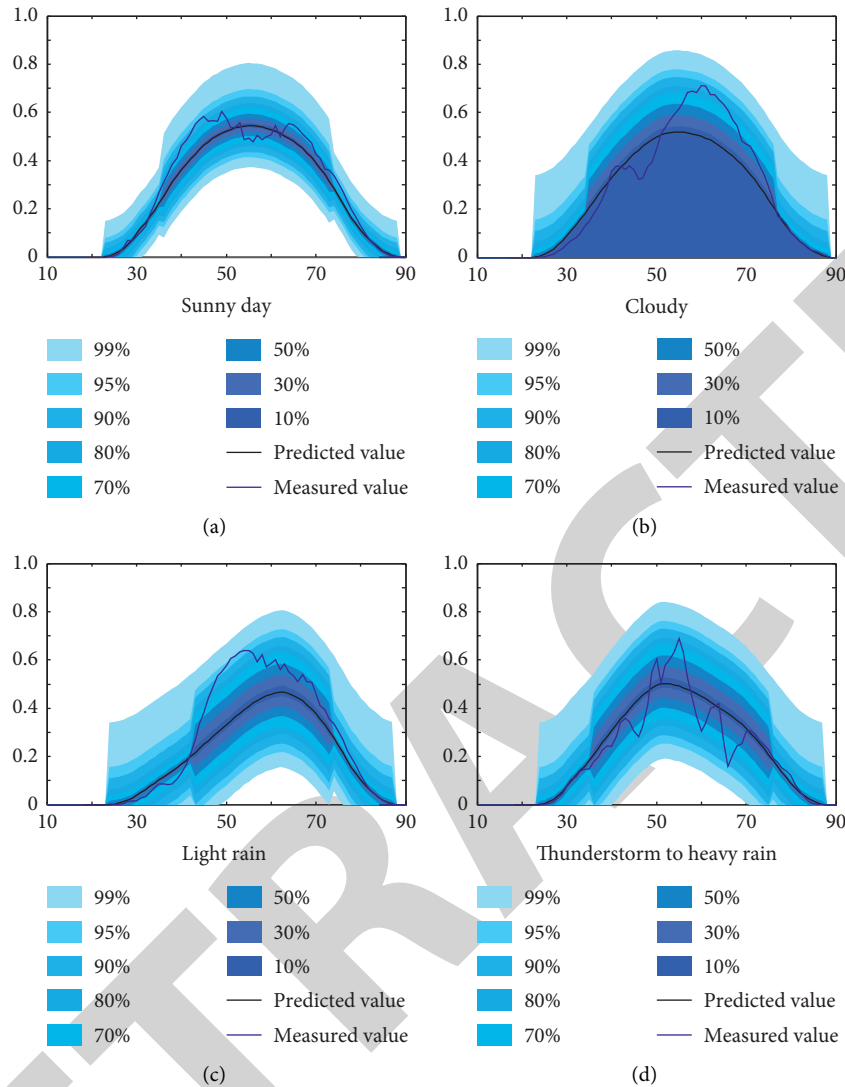


FIGURE 6: The predicted value, measured value, and GGMM confidence interval under different weather in July.

distribution. The fitting effect of normal distribution is better outside the peak value, but it is lower than the empirical value at the peak value. Class III error distributes gently outside the peak value but has prominent peak value. Therefore, when fitting Class III errors, the normal distribution and Laplace distribution are obviously deficient, and  $t$  Location-Scale is more accurate in describing the peak but obviously distorted in the nonpeak areas. The proposed general Gaussian mixture distribution has obvious advantages in describing the whole distribution. The proposed general Gaussian mixture distribution model can flexibly change the weight coefficient of each Gaussian term, so it can take into account the requirements of waist flexibility and peak value of the distribution curve and has obvious advantages in describing the short-term photovoltaic power generation output forecast error distribution.

**5.2. Applicability Analysis of Model.** In order to see whether the generalized Gaussian mixture distribution model can

perform well in different meteorological environments, the historical data of different weather type days in high temperature season: July 4th (sunny day), July 8th (cloudy day), July 17th (light rain), and July 20th (thunderstorm to heavy rain) in 2017, are selected to test the applicability of the model. Using the cluster analysis method in Section 3.3, sunny days are classified as Class I generalized weather, and cloudy, light rain and thunderstorm to heavy rain are classified as Class II generalized weather. The data are counted once every 15 minutes, and the time series points with intervals of (10, 90) are selected for analysis. The model test results are shown in Figure 6.

Figure 6 shows the predicted values, measured values, and confidence interval bands of errors of photovoltaic power generation in four different weather conditions. It can be seen from the figure that the error band width of the same confidence level is different in different weather, and the error band is the narrowest in sunny days, and the worse the weather, the wider the error band. This shows that the forecast error of photovoltaic power generation is small in

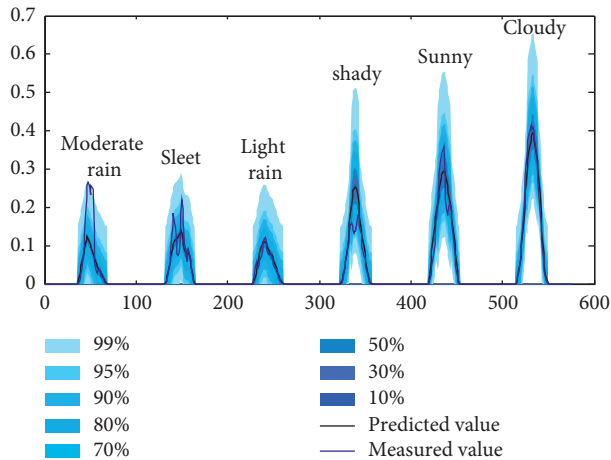


FIGURE 7: The predicted value, measured value, and GGMM confidence interval under different weather in November.

sunny days, and the probability of increasing the forecast error of photovoltaic power generation is greater with the deterioration of weather, which is consistent with the actual situation. In Class II and Class I weather, the difference between measured and predicted values is mainly concentrated at the peak value, while the measured curve at the waist is in good agreement with the predicted curve. This is because the peak belongs to the large error area, and the waist and bottom output belong to the small error area. Even so, the measured output at the peak is within the confidence interval of 95% of the predicted power.

In order to test the applicability of the model in low temperature season, the predicted, measured, and meteorological data of November 13, 14, 15, 16, 18, and 19, 2017 are selected in Figure 7 to test the applicability of the model to ambient temperature.

The forecast days selected in Figure 7 belong to Class I generalized weather. Similar to the test results in Figure 6, the measured values at the peak value deviate from the predicted values to a higher degree than those at the waist and bottom, but the measured values are all within the confidence interval of 95%, which shows that the model is very sensitive to the output value with large fluctuation.

To sum up, under different weather types, ambient temperatures, predicted output amplitude, and step size, the proposed general Gaussian mixture model can accurately describe the distribution of short-term PV power output forecast error, and the model has strong applicability. In addition, according to the weather conditions on the forecast day, the model can give the error bands under different confidence levels of PV short-term forecast power in advance.

## 6. Conclusion

Accurate description of wind and solar output uncertainty is the basis of establishing stochastic optimal dispatching model of power system with wind and solar power sources. In order to describe the short-term forecast error of photovoltaic power generation relatively accurately, a short-

term forecast error model of photovoltaic power generation output considering meteorological factors and numerical characteristics is established in this paper, and a general Gaussian mixture model is proposed to describe the short-term forecast error of photovoltaic power generation. The model considers the influence of different meteorological conditions on the forecast error and combines numerical characteristics for analysis. Finally, taking the photovoltaic power generation system in Brussels area as an example, the effectiveness of this method is verified, and the main conclusions are as follows:

- (1) The short-term PV power forecast error is affected by three weather factors: weather type, temperature difference, and maximum temperature, and is also related to the output amplitude and climbing power at the predicted time
- (2) The general Gaussian mixture model proposed in this paper can flexibly change the weight coefficient of each Gaussian probability density, so that it can take into account the requirements of waist flexibility and peak value of distribution curve at the same time, and has obvious advantages in describing the forecast error distribution of short-term photovoltaic power generation output

In this paper, the analysis of the problem is limited by the acquisition of meteorological data. If more detailed and accurate meteorological data are obtained in the future, we can further analyze the influence of meteorological factors on the forecast error at every moment in the day and establish a more comprehensive error model in order to narrow the confidence interval and obtain more accurate results.

## Data Availability

The data used to support the findings of this study are included within the article.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References

- [1] H. Wang, L. Ge, H. Li, and F. Chi, "A review on characteristic analysis and prediction method of distributed PV," *Electric Power Construction*, vol. 38, no. 07, pp. 1–9, 2017, in Chinese.
- [2] J. Wang, W. Wang, and H. Chen, "Prediction of photovoltaic power generation based on regression-Markov chain," *Electrical Measurement and Instrumentation*, vol. 56, no. 1, pp. 76–81, 2019.
- [3] Z. Wang, L. He, X. Cheng, and J. He, "Method for short-term photovoltaic generation power prediction base on weather patterns," in *Proceedings of the China International Conference on Electricity Distribution*, pp. 213–215, IEEE, Shenzhen, China, September 2014.
- [4] P. Bacher, H. Madsen, and H. Nielsen, "Online short-term solar power forecasting," *Solar Energy*, vol. 83, no. 10, pp. 1772–1783, 2009.



## Retraction

# Retracted: Discontinuous Track Recognition System Based on PolyLaneNet for Darwin-op2 Robot

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] X. Wu, S. Lv, X. Wang et al., "Discontinuous Track Recognition System Based on PolyLaneNet for Darwin-op2 Robot," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 5431886, 10 pages, 2022.

## Research Article

# Discontinuous Track Recognition System Based on PolyLaneNet for Darwin-op2 Robot

**Xi-Bao Wu, Si-Chuan Lv, Xiao-Hao Wang, Shi-Zhuo Zhang, Qiong Liu, Yi-Qun Wang ,**  
**and Wen-bai Chen **

*School of Automation, Beijing Information Science and Technology University, Beijing, China*

Correspondence should be addressed to Wen-bai Chen; [chenwb03@126.com](mailto:chenwb03@126.com)

Received 27 October 2021; Accepted 7 December 2021; Published 1 February 2022

Academic Editor: Daqing Gong

Copyright © 2022 Xi-Bao Wu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper proposes and demonstrates a single-line discontinuous track recognition system by associating the track recognition problem of a humanoid robot with the lane detection problem. The proposal enables the robot to achieve stable running on the single-line discontinuous track. The system consists of two parts: the robot end and the graphics computing end. The robot end is responsible for collecting track information and the graphics computing end is responsible for high-performance computing. These two parts use the TCP for communication. The graphics computing side uses PolyLaneNet lane detection algorithm to train the track image captured from the first perspective of the darwin-op2 robot as the data set. In the inference, the robot end sends the collected tracking images to the graphics calculation end and uses the graphics processor to accelerate the calculation. After obtaining the motion vector, it is transmitted back to the robot end. The robot end parses the motion vector to obtain the motion information of the robot so that the robot can achieve stable running on the single-line discontinuous track. The proposed system realizes the direct recognition of the first perspective image of the robot and avoids the problems of poor stability, inability of identifying curves and discontinuous lines, and other problems in the traditional line detection method. At the same time, this system adopts the method of cooperative work between the PC side and the robot by deploying the algorithm with high computational requirements on the PC side. The data transmission is carried out by stable TCP communication, which makes it possible for the robot equipped with weak computational controllers to use deep-learning-related algorithms. It also provides ideas and solutions for deploying deep-learning-related algorithms on similar low computational robots.

## 1. Introduction

The humanoid robot track and field project mainly requires the robot to reach the end point in the shortest time through various trajectories without human control. The evaluation criteria mainly include the passing time of the robot, the correct movement along the trajectory direction, and the recognition effect of obstacles on the trajectory. Common track and field runways are single continuous line and curve, double continuous line and curve, single discontinuous line and curve, double discontinuous line and curve, etc. Some tracks also include up and down slopes, up and down stairs, obstacles, crossroads, and other elements, which greatly tests the comprehensive technical level of robots and their teams. Our design focuses on identifying

single continuous and discontinuous trajectories captured by first-person robots. In essence, the track recognition of humanoid robot in track and field competition is mainly to identify one or several lines in the picture captured by the first person of the robot. These lines have different colours, continuous state, intersection, and other characteristics, which play a guiding role in the movement of robots, similar to the track line of human track and field competition and the lane line of vehicles driving on the road. The robot uses various algorithms to identify the track and obtains the offset position and guiding direction of different forms of track. Then, the postprocessing analysis is carried out on the offset position and orientation direction of the trajectory to determine the motion direction of the robot at the next moment, using a fixed step to modify the

lateral offset of the movement or directly providing a direction vector.

There are many kinds of track recognition algorithms for humanoid robot track and field competitions. The simplest algorithm is the recognition method based on photoelectric sensor, which uses the colour difference between the track line colour and the track bottom colour to identify. Some of the complex algorithms are various visual algorithms based on cameras. The camera deployed on the robot captures the first person or non-first-person real-time images of the robot for recognition, including traditional recognition algorithms such as pixel-by-pixel comparison method, Hough [1] transform method, connected domain recognition method, line segments angles computation [2], and deep learning recognition algorithms [3–14] such as deep convolutional neural network recognition algorithm. In the track and field competition of humanoid robot, single-line or double-line noncontinuous track is often more difficult to identify than continuous track. In order to solve the problem of rapid and accurate recognition of single-line noncontinuous track by robots, this paper comprehensively examines various visual recognition algorithms and nonvisual recognition algorithms, combined with the lane line detection algorithm PolyLaneNet [15]; a deep learning algorithm is deployed on the humanoid robot with low computing power, so that it can accurately identify the track line in the track and field competition of humanoid robot, so as to achieve the effect of high-level competition.

PolyLaneNet algorithm constructs the recognition problem of lines into the task of polynomial formula analysis of lines. Common lanes or racetracks can be resolved as an  $N$ -order polynomial ( $N$  is the natural number). PolyLaneNet first puts the image into the deep convolutional neural network (one of EfficientNet [16, 17], ResNet34 [18], ResNet50, and ResNet101) for feature extraction and then the extracted features are postprocessed to obtain the polynomial formula that can accurately describe the line. It ensures high recognition accuracy and has high computing speed. The PolyLaneNet algorithm was originally used to solve the lane line detection problem. Because the lane line recognition problem in humanoid robot track and field competition is similar to the lane line detection problem and even in most cases it is simpler than the lane line problem (reducing the occlusion of the lane line by the side car, there are many lanes, there are many types of lane lines, and the lane line is thicker and more obvious than the lane line and it is easier to identify), this paper will transplant the PolyLaneNet algorithm with excellent performance in the lane line detection problem to the track recognition problem of humanoid robot track and field competition.

The single-line noncontinuous track recognition system designed in this paper consists of two parts: the robot end and the graphics computing end. The graphics computing end is composed of a computer or GPU (graphics processing unit) server containing an independent graphics computing card. After the robot camera captures the real-time picture, the picture is transmitted to the graphics computing end through TCP (Transmission Control Protocol) communication. The graphics computing end performs

PolyLaneNet algorithm reasoning on the real-time picture received by TCP communication and processes the reasoning results to obtain the direction vector that the robot should move next. The direction vector is transmitted back to the robot through TCP communication. The robot receives the direction vector and controls the steering gear to move. The average speed and recognition stability of darwin-op2 robot in this scene have been significantly improved. Specifically, the darwin-op2 robot runs stably on the track and field track at its own 0.5 times speed (actual 0.12 m/s) and 0.7 times speed (actual 0.15 m/s), which meets the running requirements of the robot on the track and field track. And the robot performed well even at full speed, only occasionally running out of the field. Good performance has been achieved in many experiments and controls.

## 2. Relevant Algorithms

*2.1. Track Line Recognition Algorithms.* In the humanoid robot track and field race, the teams used various methods to identify the track line. The following will introduce several common recognition methods in humanoid robot track and field competition.

To begin with, the first recognition method is based on a photoelectric sensor. The photoelectric sensor has a light intensity receiving terminal, which can convert the received light intensity into an electrical signal with analog value. Due to the difference between the colour of the track line and the background colour of the track in the race, the reflected light intensity is different and the analog value converted into electrical signal is also different. The recognition algorithm based on photoelectric sensor uses the above principle to recognize the track line and distinguishes the track line from the track background by distinguishing different simulation value intervals. The disadvantage of this recognition method is that it is easily affected by the external light intensity environment, the threshold changes greatly, and the generalization ability of different light conditions is weak. It needs to adjust the sensor threshold before each race. And because of the characteristics of the sensor itself, it cannot support the rapid movement of the robot and cannot recognize the discontinuous track.

The second pixel-by-pixel comparison method is based on camera. Firstly, the image captured by the camera is preprocessed by Gaussian filtering, binarization, expansion, and corrosion to get a binarization image, where the track line is white and the background is black or the track line is black and the background is white. Then, the pixels of the binary image are traversed according to a certain rule and the pixel position of the track line in the picture is obtained. The disadvantage of this algorithm is that it is easy to form false recognition to the noise or large area colour block after image preprocessing; also, the recognition effect of discontinuous track is poor.

Furthermore, the connected domain recognition algorithm based on camera is used frequently. Firstly, the image captured by the camera is preprocessed by Gaussian filtering, binarization, expansion, and corrosion to get a binary image. Then, the connected domain of the binary image is identified

and the largest connected domain is the track line. The lateral offset, inclination angle, and other information of the track line can be obtained by calculating the connected domain. The algorithm still cannot eliminate the interference of large area colour block and the recognition effect of discontinuous track is also poor.

The Hough transform algorithm based on camera is also one of recognition methods. The image is preprocessed by Gaussian filtering, edge detection, and binarization. Then, the image is mapped from the Cartesian coordinate system to the polar coordinate system. The algorithm takes the local maximum value, sets the threshold, and filters the noise for each point in the polar coordinate system. Finally, the points in polar coordinate system are transformed back to Cartesian coordinate system, which is the straight line in the original image recognized by the algorithm. This algorithm can identify the interference of discontinuous track, immune noise, and large-area colour blocks and has strong recognition ability for straight lines. However, it is easy to form false detection for long strip interference and has weak recognition ability for curves.

**2.2. Lane Detection Algorithms.** The common methods of track line identification described above all have defects in different aspects. In this paper, we notice that the recognition of track lines in humanoid robot track and field events is essentially to find or fit one or more long lines in an ROI (region of interest) region of the image, which is used to guide the robot to a certain direction. The problem of lane detection in the field of automatic driving is essentially to find or fit the lane in an ROI region of the image, which is used to guide the car to a certain direction. At the same time, the lane detection is similar to the track recognition; there are continuous line and discontinuous line, different degrees of noise interference and occlusion, and real-time problems. Therefore, this paper considers that the problem of track line recognition in humanoid robot track and field events and the problem of lane line detection in automatic driving are essentially the same kind of problems. They are all used to identify one or more lines in the ROI region to guide the direction of movement, which can be solved by the same or similar methods. The problem of lane detection takes the colour picture as the input and the image segmentation examples, some points, or parameters as the output. This paper summarizes some common lane detection methods as follows.

The traditional lane detection algorithm is mainly based on Hough transform and the detection method is the same as the Hough transform detection method mentioned above. In recent years, with the rapid development of deep learning and the continuous success of convolutional neural network in many computers vision tasks, various research teams gradually move the solution of lane detection problem closer to the deep learning method. At present, several commonly used lane detection methods based on deep learning mainly include detection method, grid method, polynomial regression method, and anchor-based detection method.

Detection methods mainly include Lanenet [19], Scnn [20], Sad, and so on. Lanenet algorithm is a classic algorithm and it is widely used in industry. It belongs to instance segmentation algorithm, but its disadvantage is that it relies on clustering and the detection result is unstable. Scnn algorithm is a semantic segmentation algorithm, which uses the method similar to RNN to learn the spatiotemporal information. It defines K lines in advance, makes K classification prediction for each pixel, and finally combines the points with the same results to get the lane lines to be detected, but the disadvantage is that the recognition speed is slow. Sad algorithm is also a segmentation algorithm; its structure is lighter than Scnn algorithm, so the recognition speed is faster than the Scnn algorithm, but the recognition accuracy is relatively low.

The main idea of the gridding method is to turn the pixel problem of the image into a cell problem, which is similar to the idea of downsampling. The recognition speed is faster, but because it is equivalent to reducing the resolution of the image, the recognition accuracy of the algorithm is low.

Polynomial regression method is to fit the lane line into an n-order polynomial (n is a natural number). The speed of recognition is faster, but the disadvantage of the polynomial regression method is that it relies on a priori and has low flexibility in complex scenes

The detection method based on anchor point is similar to the target detection method. Some lines with various angles are defined in advance and their offsets are learned. This detection method also depends on a priori and has low flexibility in complex scenes.

### 3. Algorithm Selection

**3.1. The Original Algorithm.** Before carrying the track line recognition algorithm used in this design, the track line recognition algorithm used by our team is the pixel-by-pixel comparison method based on the camera mentioned above.

First, a camera is used to capture an image of the robot from the first view. Then, the image is transformed from a three-channel colour image to a single-channel grey image and the grey image is binarized according to a certain threshold (the threshold needs to be adjusted manually according to the colour and illumination of the competition venue), so as to change the colour of the track into black and the colour of the track line into white. The transformation process is shown in Figure 1.

Then, the binary image is divided into two regions (the height range of the region is adjustable) and the two regions are compared pixel by pixel to find the track line. After the lateral offsets of the track lines in the upper and lower regions are obtained, the weighted sum of the two lateral offsets and the offset are added based on some decision-making methods. The final offset is the lateral offset that the robot needs to move forward at the next moment.

The specific process of pixel-by-pixel comparison is to traverse the pixels of each region from top to bottom. For each row of pixels, traverse from the middle to both sides. When three consecutive white pixels are encountered, it is the track line that meets white (after binarization). The

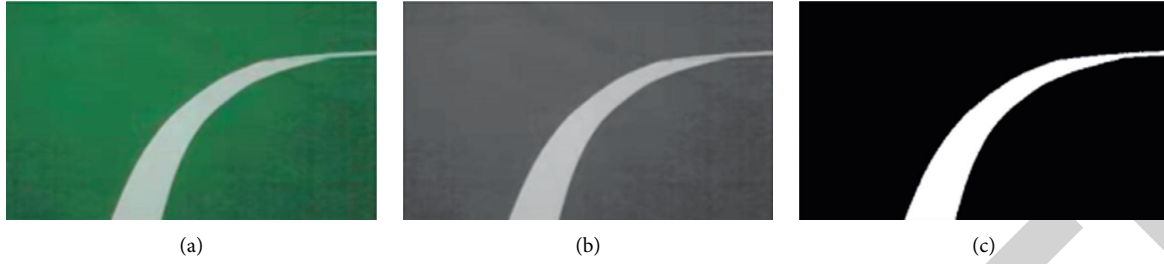


FIGURE 1: Image processing by pixel comparison method: original image (a), grey image (b), and binary image (c).

abscissa position of the pixel is recorded and stored in the array. After the end of line-by-line traversal, the abscissa of the white track line found in each line and stored in the array is averaged and then the abscissa of the middle point in the image is subtracted to get the average lateral offset of the current area track. If there is no track in the current area, the lateral offset is 0. The time complexity of this method is  $O(n^2)$ .

The pixel-by-pixel comparison method based on camera has simple logic and short time consumption, but it compares the pixels, only identifies the local, and then combines the results of local recognition to form a global decision. The disadvantage of local recognition is that it is easy to be interfered by noise and large area colour block (as shown in Figure 2). Once there is interference in the picture which is far away from the real track line, the final global decision will be inaccurate. Because of its line-by-line recognition of pixels and small track area (generally speaking, the camera only captures a small part of the track in front of the robot), once it encounters a discontinuous track line, it will not find the track line.

**3.2. The PolyLaneNet Algorithm.** In order to reduce or eliminate the noise and large area colour block interference and realize the recognition of discontinuous track, we need an algorithm to recognize the global information and increase the area of the track captured by the robot first person picture. Common algorithms such as Hoff transform map Cartesian coordinate system to polar coordinate system, reorganize all local information, then carry out local recognition in polar coordinate system and use global information in disguised way, eliminate noise and nonlong bar colour block interference, and realize recognition of discontinuous linear track, but they cannot recognize curve track better. Convolution neural network can recognize local information and global information by using many convolutions check images for several times and extract higher dimension features, which can fully satisfy the interference of noise and large area colour block and can recognize not only discontinuous linear track, but also discontinuous curve track, The ability of anti-jamming is enhanced and the ability of track line recognition is improved. The disadvantage is that the calculation is large and a large amount of data is needed to train the model. After considering several methods mentioned above, this paper finally decides to use the method of deep learning (deep convolution neural network) to identify the track line. The method of deploying the model on PC (personal computer)

and communicating with robot through wireless LAN (local area network) solves the problem that the computer human controller cannot bear the convolution neural network calculation.

The previous paper lists several different principles and effects of lane detection methods based on deep learning. Compared with lane detection in automatic driving, the scene of track line recognition in humanoid robot track and field race is simpler (as shown in Figure 3), but it requires higher real-time performance. Therefore, this paper selects polynomial regression method, which has faster recognition speed and higher recognition accuracy, but higher requirements for scene environment. The PolyLaneNet (lane estimation via deep polynomial regression) algorithm proposed in 2020 is an excellent polynomial regression algorithm in lane detection. The theoretical FPS (frames per second) of the algorithm is up to 115 and the accuracy of the algorithm is up to 93% in TuSimple data set. If it is transplanted to the track line recognition of humanoid robot track and field race, its high FPS can reduce the time cost and improve the real-time performance as much as possible and the simpler scene in the track line recognition also reduces the false detection and missing detection rate of the algorithm and ensures the accuracy of the recognition. Therefore, this paper uses PolyLaneNet algorithm to identify the track line.

## 4. Algorithm Design Based on PolyLaneNet

**4.1. Principle Overview.** In practical application, there are many problems in the recognition of lane line or track line in track and field competition, such as occlusion, wear, discontinuous line, and so on. The shape of the line is thin and long, which leads to sparse supervision signal, which is difficult to detect. At the same time, the task requires high real-time. Based on convolutional neural network, the PolyLaneNet algorithm used in this paper uses ResNet series network or EfficientNet network as backbones to extract features from images. The extracted features pass through a layer of full connection layer to get several one-dimensional vectors; each vector contains the information of the highest point, the lowest point of a line in the graph confidence information and polynomial coefficient information. The algorithm structure is shown in Figure 4.

PolyLaneNet expects as input images taken from a forward-looking vertical camera, and outputs, for each image,  $M_{\max}$  lane marking candidates (represented as polynomials), as well as the vertical position  $h$  of the horizon line, which helps to define the upper limit of the lane



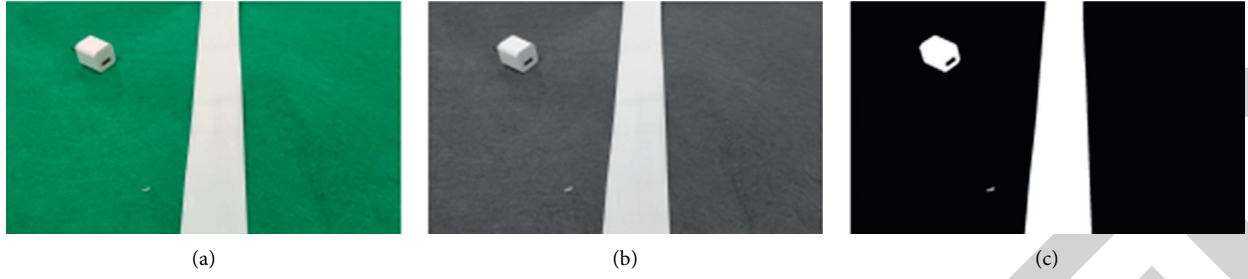


FIGURE 2: Image processing with noise and colour block: original image (a), grey image (b), and binary image (c).



FIGURE 3: Lane detection pending image (a) and track detection pending image (b).

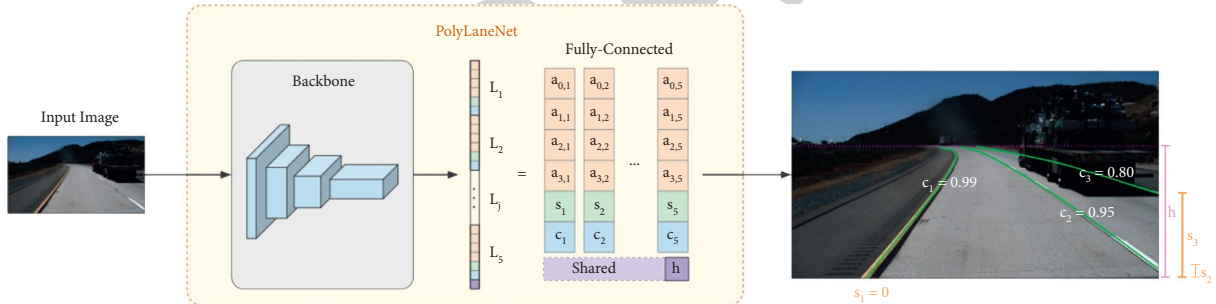


FIGURE 4: PolyLaneNet structure diagram.

markings. The architecture of PolyLaneNet consists of a backbone network (for feature extraction) appended with a fully connected layer with  $M_{\max} + 1$  outputs, being the output  $1, \dots, M_{\max}$  for lane marking prediction and the output  $M_{\max} + 1$  for  $h$ . PolyLaneNet adopts a polynomial representation for the lane markings instead of a set of points. Therefore, for each output  $j, j = 1, \dots, M_{\max}$ , the model estimates the coefficients  $p_j = \{a_{k,j}\}_{k=0}^K = 0$  representing the polynomial, shown as follows:

$$p_j(y) = \sum_{k=0}^K a_{k,j} y^k, \quad (1)$$

where  $K$  is a parameter that defines the order of the polynomial. As illustrated in Figure 4, the polynomials have restricted domain: the height of the image. Besides the coefficients, the PolyLaneNet estimates, for each lane marking  $j$ , the vertical offset  $s_j$ , and the prediction confidence score  $c_j \in [0, 1]$ . In summary, the PolyLaneNet model can be expressed as follows:

$$f(I; \theta) = \left( \{p_j, s_j, c_j\}_{j=1}^{M_{\max}}, h \right), \quad (2)$$

where  $I$  is the input image and  $\theta$  is the model parameters. At inference time, as illustrated in Figure 4, only the lane marking candidates whose confidence score is greater than or equal TO a threshold are considered as detected.

**4.2. Backbone Process.** The original meaning of backbone is human backbone and then extended to the meaning of pillar, core, and so on. Backbone represents the foundation of the whole model in deep learning, because the subsequent tasks such as classification, detection, and generation are based on the extracted features, so in the visual field, backbone is the process of image feature extraction. In the field of machine vision, depth convolution neural network is mainly used to extract image features. Convolution neural network uses convolution check image of specified size to convolute and obtains image texture and other features and then uses these features for various postprocessing. Convolution neural



network uses back propagation, gradient descent, and other methods to learn and update the parameters of convolution kernel and the weight parameters of convolution kernel do not need to be adjusted manually. Due to the strong non-linear modelling ability of neural network, convolutional neural network can solve most problems in many application scenarios. The basic structure of convolutional neural network is shown in Figure 5.

PolyLaneNet algorithm takes ResNet series network or EfficientNet series network as the backbone of the model for feature extraction. ResNet network and EfficientNet network belong to deep convolution neural network. The ResNet series network realizes the deeper development of the network through the “residual unit,” which enables the network to recognize more advanced features. The EfficientNet series network comprehensively improves the depth, width, and resolution of the input image of the network and uses the network structure search method to search the network structure, so as to obtain the results of low resource consumption, high efficiency, and low-cost network with high recognition accuracy.

**4.3. Posttreatment Process.** The input image is transformed into a one-dimensional vector with the length of  $(6 * n + 1)$  after feature extraction by the depth convolution neural network in the backbone. We transform it into a two-dimensional feature of  $7 * n$  (7 one-dimensional vectors) for output, which contains the information of  $N$  lines. The information of each line includes four polynomial coefficients, one ordinate of the lowest point, one ordinate of the highest point, and one confidence level. After several steps of postprocessing, we need to use an array of points to represent the recognized line.

First of all, we filter out the vector that does not recognize the line (the vector with the confidence of 0) and leave the vector that recognizes the line (because this system does single track identification, only the vector with the highest reliability is retained). Then, we traverse these vectors, extract the vertical axis coordinates of the highest point and the lowest point of each line, construct an array of 100 based on the two vertical coordinates, and divide its height into 100 equal parts. Then, we fit the line according to the four polynomial coefficients and the array with a length of 100 and combine the abscissa calculated with the ordinate divided into 100 parts to form 100 coordinates, which represent the specific shape of the line we fit.

Finally, the decision-making process of the fitted line is carried out. We take the highest point, the middle point, and the lowest point to calculate the weighted difference and get the direction vector which needs to control the robot at the next moment and send it to the robot through TCP communication.

The whole postprocessing flow chart is shown in Figure 6.

## 5. Realization of the Discontinuous Track System

**5.1. System Flow.** Because the computing power of cm740, the core control board of darwin-op2 robot, is not enough to

support PolyLaneNet algorithm for reasoning, the gtx1650 super graphics card on PC is used for model reasoning. The information transmission between robot and PC is supported by wireless LAN.

The single-line discontinuous track recognition system of the whole robot consists of discontinuous track, robot, PC terminal, and a local area network, as shown in Figure 7.

Because the communication between robot and PC requires high stability and UDP (User Datagram Protocol) communication has the problem of packet loss, we choose more stable TCP communication instead of UDP communication.

The whole operation process of the system is as follows:

- (i) The robot and PC handshake in LAN to establish TCP communication.
- (ii) The robot camera captures the single track picture from the first perspective, packages the picture into a socket package, and transmits it to the PC through TCP communication.
- (iii) After receiving the socket packet, the PC transcodes the data into pictures.
- (iv) On the PC side, the transcoded image is put into the PolyLaneNet algorithm for forward reasoning and  $N$  one-dimensional vectors are obtained after feature extraction by backbone.
- (v) The PC end carries on the postprocessing to the  $N$  one-dimensional vectors obtained from the forward reasoning, fits the array which can represent the track line, and extracts the coordinates of the highest point, the middle point, and the lowest point of the track line.
- (vi) The PC end calculates the coordinates of the highest point, middle point, and lowest point of the extracted track line, obtains the direction vector of the track line (that is, the direction vector that the robot needs to move at the next moment), and packages it into a socket package, which is sent to the robot through TCP communication.
- (vii) After receiving the socket packet, the robot transcodes the data into an array vector and transmits the value of the vector to the API that controls the robot to run in the specified direction.
- (viii) At the next moment, the robot camera will continue to capture the single track picture from the first perspective and repeat steps 2–7.

The operation flow chart of the system is shown in Figure 8.

**5.2. Data Set.** In this paper, 3141 images with common floor as the track and 5 cm purple line as the track line are collected and the RGB three channels are rearranged and combined and a data set of 18846 images is obtained. The data set is shown in Figure 9.

On the PC side, the image is first scaled to  $640 * 360$  for saving and the data set is shown in Figure 9. Then, the image is annotated with “Label Me” annotation software. Each

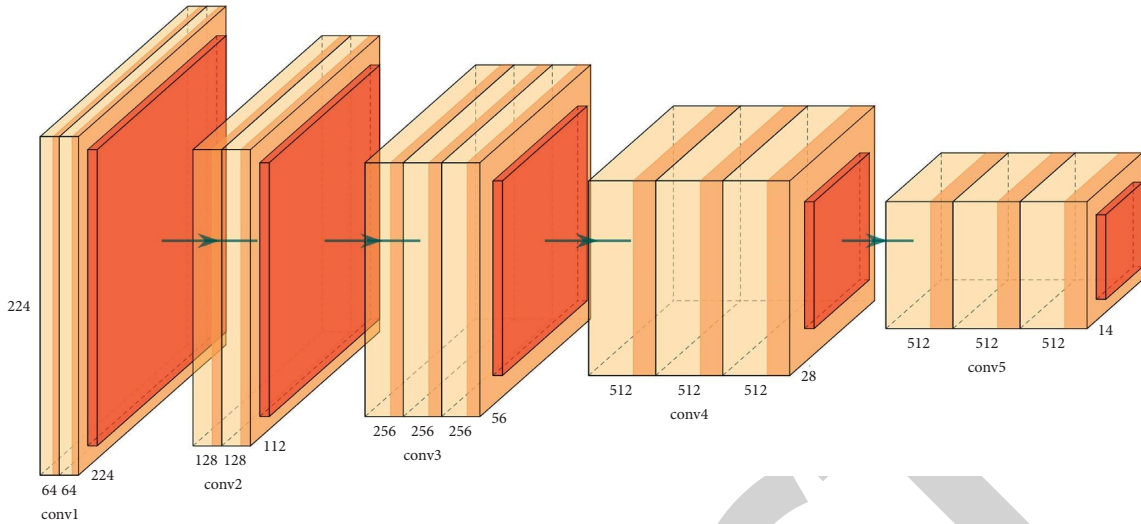


FIGURE 5: Basic structure of convolutional neural network.

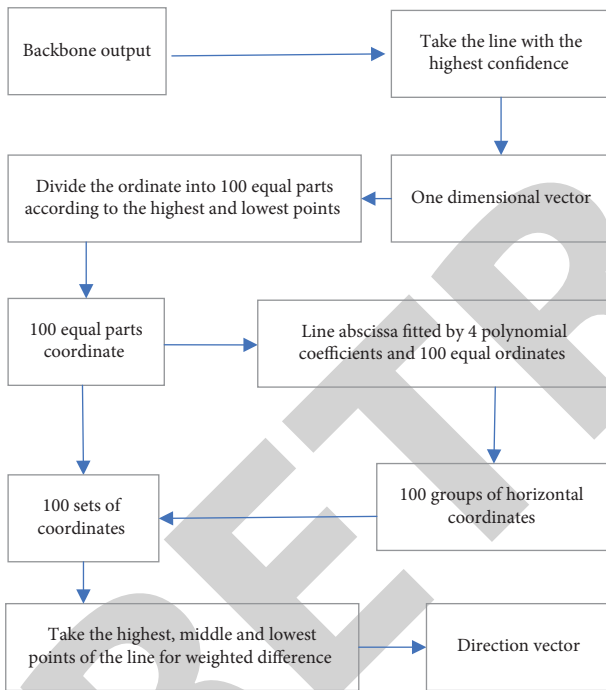


FIGURE 6: Postprocessing flow.

picture is marked with 7 points to represent the track line, as shown in Figure 10.

## 6. Results and Optimization

**6.1. Preliminary Results.** The system is tested on a green track and 5 cm wide white track. In order to ensure the stability of recognition, the half speed of the normal speed (0.24 m/s) of the robot is used to test and the average speed is about 0.11 m/s. The average speed is about 0.14 m/s and the recognition stability is poor. Finally, the normal speed is used to test, the average speed is about 0.15 m/s, and the

recognition stability is very poor. The FPS of the system in three cases is always 6–9. The experimental results are summarized in Table 1.

**6.2. Increasing the FPS of the System.** According to the preliminary results, it can be seen that the FPS of the system is very low, which leads to the slow reaction speed of the robot, resulting in its frequent off track. This problem can be alleviated and solved by reducing the movement speed of the robot and the time consumption in the system program. The generalization ability of the system is poor and the robot can only recognize the green track and white track line. If the colour of the track or track line is changed, the robot cannot recognize it. This problem can be solved by using more data sets for training and normalizing the colour of the screen.

Because the ultimate goal of humanoid robot track and field competition is to reach the destination with the fastest speed and reducing the robot’s movement speed can enhance the stability, but it is contrary to the ultimate goal, this method can cure the symptoms but not the root cause. Therefore, it is necessary to improve the system FPS by reducing the time consumption of the system program. Through the separate detection of the time consumption of each process of the system, it is found that the most time-consuming process is the TCP communication transmission picture part, the postprocessing part of the PC algorithm, and the control part of the robot movement. Because the part that controls the robot’s movement calls the underlying API directly, it cannot be optimized, so this paper attempts to optimize the TCP communication part and the algorithm postprocessing part.

In this paper, the main details of the screen are not lost and the recognition accuracy rate is not affected. Firstly, the image is reduced from 320 \* 240 to 40 \* 22 resolution and then packaged into socket packet for

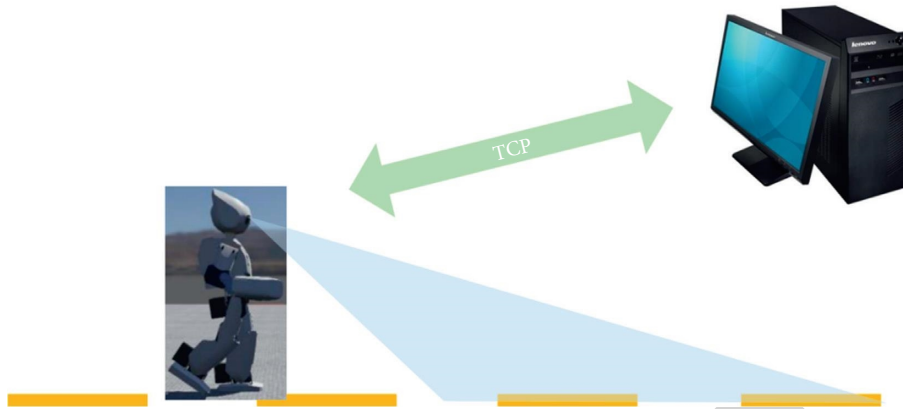


FIGURE 7: Schematic diagram of track and field competition between the robot and PC.

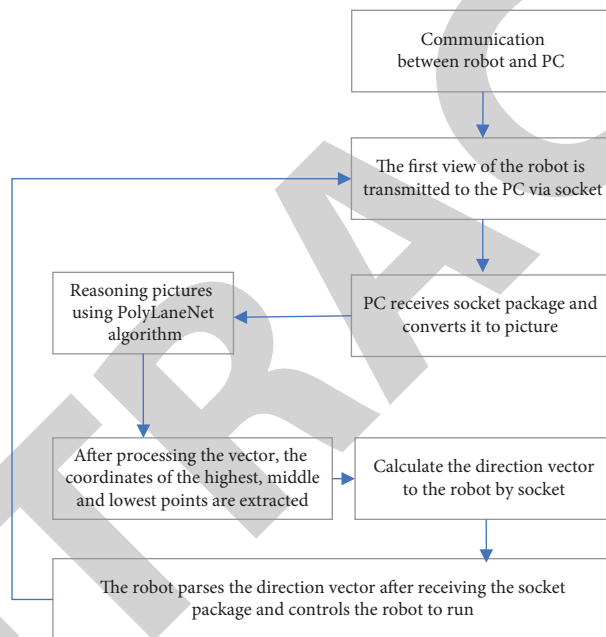


FIGURE 8: System flow chart.

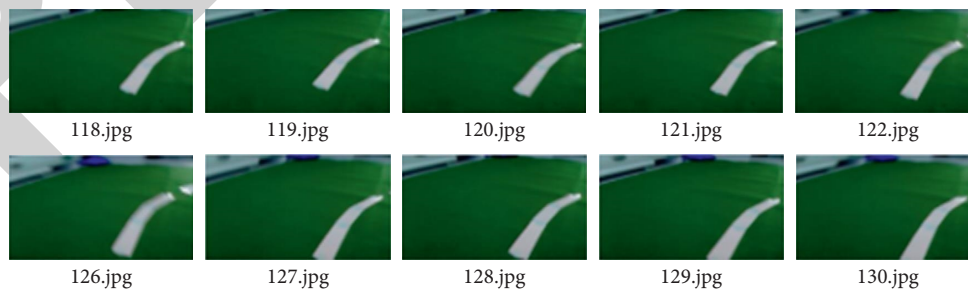


FIGURE 9: Filtered and scaled data set.

transmission. This greatly reduces the amount of data when TCP communication transmits the picture, thus greatly shortening the transmission time. After the image is compressed, the proportion can be reduced by half.

While the main information of the image is retained, the data quantity changes to 0.01146 times of the original, which makes the time consuming of transmission picture greatly reduced.

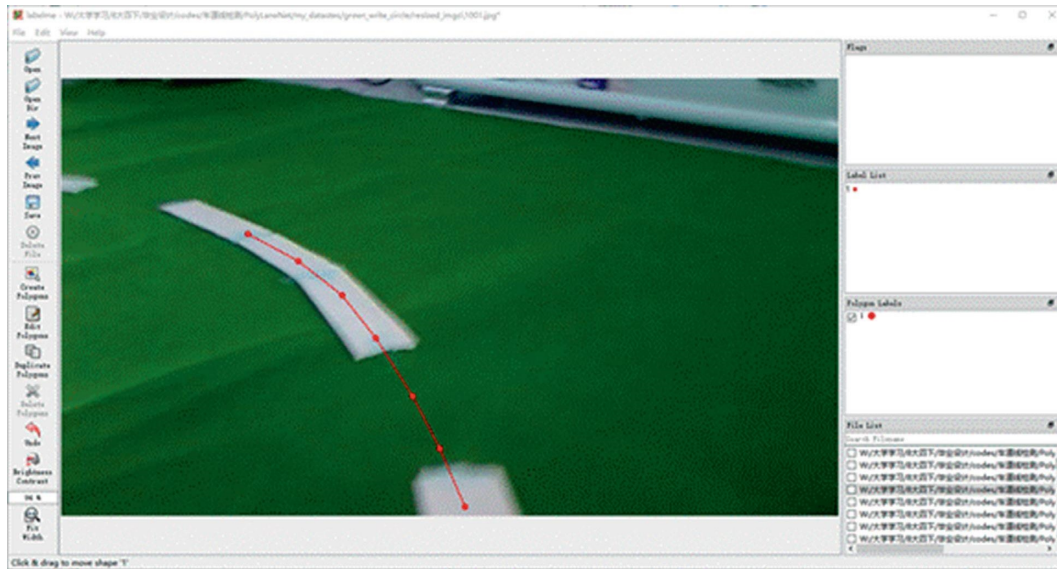


FIGURE 10: Labeling data sets.

TABLE 1: Summary of experimental results.

Theoretical velocity	Actual average speed	Off track times	FPS
0.5 times speed	0.11 m/s	1	6–9
0.7 times speed	0.14 m/s	3	6–9
1 time speed	0.15 m/s	6	6–9

TABLE 2: Summary of experimental results after system improvement.

Theoretical velocity	Actual average speed	Off-track times	FPS
0.5 times speed	0.11 m/s	0	12–15
0.7 times speed	0.14 m/s	0	12–15
1 time speed	0.15 m/s	1	12–15

The results after system optimization are shown in Table 2.

It can be seen that the average speed, recognition stability, and system FPS of the robot in this scene have been significantly improved, which meets the requirements of the robot running on the track and field track.

## 7. Conclusions

This design uses the method based on deep learning (deep convolution neural network) to detect the single line, curve, continuous, and discontinuous track line in humanoid robot track and field race. It is found that the essence of the track line detection problem in humanoid robot track and field race is the same as that in automatic driving. Therefore, the solution of lane detection in automatic driving is applied to the problem of track detection in humanoid robot track and field events. The high-FPS and high-precision PolyLaneNet model is selected to detect the single track continuous and

discontinuous track lines in track and field events. Thanks to the strong nonlinear modelling ability of deep neural network, the system can recognize the first-person image of robot directly and avoid all kinds of problems in traditional line detection methods (poor stability, unable to recognize curves, discontinuous lines, etc.). However, due to the limited computing power of the cm740 controller of darwin-op2 robot, it is unable to directly carry the deep neural network. Therefore, this system uses the method of working with the robot on the PC side, deploys the PolyLaneNet algorithm with high computing power demand on the PC side, uses the LAN to connect the PC side and the robot, and transmits data through stable TCP communication. It makes it possible for darwin-op2 robot equipped with weak computing power controller to use deep learning correlation algorithm and also provides ideas and solutions for deploying deep learning correlation algorithm on similar low computing power robots.

Subsequently, we will solve the problem of noise in the result vector obtained by the PolyLaneNet algorithm of the picture captured by the camera due to strong jitter during robot running and we can alleviate and solve the problem by using Kalman filter and other methods for the result vector obtained by the PolyLaneNet algorithm. Using dynamic information, Kalman filter can not only filter out the noise caused by picture jitter, but also predict the guidance direction of the track line at the next moment. Theoretically, it can alleviate the noise problem caused by strong picture jitter to a certain extent.

## Data Availability

No data were used to support this study.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## *Retraction*

# **Retracted: Modeling and Optimization of Water Distribution in Mineral Processing considering Water Cost and Recycled Water**

### **Computational Intelligence and Neuroscience**

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### **References**

- [1] H. Wang, Q. Cui, and H. Du, "Modeling and Optimization of Water Distribution in Mineral Processing considering Water Cost and Recycled Water," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 2314788, 12 pages, 2022.



## Research Article

# Modeling and Optimization of Water Distribution in Mineral Processing considering Water Cost and Recycled Water

He Wang , Qiuyu Cui, and Hualong Du

*School of Mechanical Engineering and Automation, University of Science and Technology Liaoning, Anshan 114051, China*

Correspondence should be addressed to He Wang; wanghe@ustl.edu.cn

Received 25 October 2021; Accepted 17 January 2022; Published 1 February 2022

Academic Editor: Daqing Gong

Copyright © 2022 He Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Reducing mineral processing water costs and freshwater consumption is a challenging task in the mineral processing water distribution (MPWD). The work presented in this paper focuses on two aspects of the MPWD optimization model and the MPWD optimization method. To achieve MPWD optimization effectively, a nonlinear constrained multiobjective model is built. The problem is formulated with two objectives of minimizing the mineral processing water costs and maximizing the amount of recycled water. In this paper, an optimization method named enhancing the multiobjective artificial bee colony (EMOABC) algorithm is proposed to solve this model. The EMOABC algorithm uses four strategies to obtain the Pareto-optimal solutions and to achieve the MPWD optimal solutions. With the three benchmark functions, the EMOABC algorithm outperforms the other two widely used algorithms in solving complex multiobjective optimization problems. The EMOABC algorithm is then applied to two cases. Results have shown that the proposed algorithm has the ability to solve the MPWD optimization model. The developed model and the proposed algorithm provide decision support for the actual MPWD problem.

## 1. Introduction

Mineral processing is the process of obtaining the raw materials for smelting and consumes large amounts of water. For water distribution, the decision-makers in mineral processing plants need to make decisions concerning different objectives, such as water costs and freshwater minimization, to maximize the comprehensive benefits of mineral processing plants. The solution to the multiobjective optimization problem often results from both an optimization model and an optimization algorithm. Thus, this paper focuses on solving mineral processing water distribution (MPWD) optimization problems from these two perspectives.

Over the years, many works have examined the water distribution optimization problem and presented many mathematical models and conventional methods such as linear programming [1, 2], nonlinear programming [3], and integer linear programming [4], which have been applied to solve the problem. However, these methods suffer from severe limitations in handling discreteness, nonlinearity,

complex constraints, and local convergence [5]. To overcome the limitations of classical optimization methods, the heuristic methods have been proposed to solve water distribution optimization, such as simulated annealing (SA) algorithms [6, 7], genetic algorithms (GA) [5, 8, 9], differential evolution (DE) algorithms [10–12], particle swarm optimization (PSO) [13–15], ant colony optimization (ACO) [16–18], strength Pareto evolutionary algorithm (SPEA) [19], and shuffled frog leaping algorithm (SFLA) [20, 21]. Despite the success of these heuristic methods in certain aspects, they have not been applied to solve MPWD problems. Currently, MPWD problems are mainly based on manual scheduling. Therefore, building MPWD model and applying heuristic methods to solve the model represent an important issue of our current concern.

Heuristic methods have been widely used in various fields [22–25]. The artificial bee colony (ABC) algorithm is a relatively new heuristic method that has been proven to be an excellent competitor in multiobjective problems [26]. Several multiobjective ABC algorithms (MOABCs) have been proposed for multiobjective problems [27–29].



However, compared to the huge in-depth studies of other heuristic methods, such as nondominated sorting genetic algorithm II (NSGA-II) [30] and multiobjective particle swarm optimization (MOPSO) [31], on multiobjective problems, improving the performance of MOABCs is still a challenge.

MPWD is a complex process in which interrelated factors are multiple. In addition to meeting the total amount of mineral processing water, many factors, such as the flow of the water source, the pressure of the water source, and maximizing the use of recycled water, need to be considered. The traditional linear model is difficult to describe the relationship between these complex factors for MPWD with the above characteristics and cannot treat multiple constraints appropriately.

The purpose of this paper is to build a nonlinear and multiobjective model for MPWD and to develop an MOABC algorithm for solving the MPWD optimization problem. The MPWD optimization model has two objectives: one objective is to minimize mineral processing water costs, and the other objective is to maximize the amount of recycled water. The proposed MOABC algorithm called enhancing multiobjective artificial bee colony (EMOABC) algorithm integrated multiple strategies to improve its performance. Therefore, some of the contributions of this paper are as follows:

- (1) The developed model for MPWD is the nonlinear model that has two objective functions
- (2) The proposed EMOABC algorithm can improve the population diversity and obtain better global convergence

The rest of the paper is organized as follows: In Section 2, the MPWD optimization model is built. Section 3 first gives a review of the basic ABC algorithm and then proposes the EMOABC and presents the results of the EMOABC algorithm compared with other algorithms on three widely used benchmark functions. In Section 4, the implementation of the EMOABC on the MPWD optimization problem is presented. Section 5 outlines the conclusions.

## 2. MPWD Model

According to categories of water source, the mineral processing water can be divided into fresh water, tailings water, clean water, and in-plant water, and they are from rivers and lakes, tailings reservoirs, wastewater treatment plants, and in-plant recycled systems, respectively (Figure 1). The cost of using the tailings water, the clean water, and in-plant water is relatively high for proper treatment and long-distance transmission before entering into the mineral processing plant. Although the cost of using fresh water is very low due to no treatment required, it is important to minimize the use of fresh water for environmental reasons. Mineral processing water becomes wastewater after mineral processing. The MPWD problem in this paper is to optimize two competing objective functions simultaneously, minimizing the mineral processing costs and maximizing the amount of recycled water including tailings water, clean

water, and in-plant water, while satisfying several equality and inequality constraints. Generally, the problem is formulated as follows.

**2.1. Decision Variable.** The vector of the mineral processing water is expressed as  $\mathbf{Q}$ , including the freshwater as  $\mathbf{Q}_1$ , tailings water as  $\mathbf{Q}_2$ , clean water as  $\mathbf{Q}_3$ , and in-plant water as  $\mathbf{Q}_4$ . The mineral processing water except freshwater is called the recycled water. The vector of the recycled water is expressed as  $\mathbf{Q}' = \{\mathbf{Q}'_1 | \mathbf{Q}_2 \cup \mathbf{Q}_3 \cup \mathbf{Q}_4\}$ .  $q_i$  is a flow of the  $i$ th water source;  $a$ ,  $b$ ,  $c$ , and  $n$  are positive integers, and  $n$  is the number of water sources. Decision variables are defined as follows:

$$\begin{aligned} \mathbf{Q} &= (q_1, q_2, \dots, q_n), \\ \mathbf{Q}_1 &= (q_1, q_2, \dots, q_a), \\ \mathbf{Q}_2 &= (q_{a+1}, q_{a+2}, \dots, q_b), \\ \mathbf{Q}_3 &= (q_{b+1}, q_{b+2}, \dots, q_c), \\ \mathbf{Q}_4 &= (q_{c+1}, q_{c+2}, \dots, q_n). \end{aligned} \quad (1)$$

**2.2. Objective Function.** Let  $c_i$  be the cost coefficient of the  $i$ th water source; its meaning is the total power consumption required to lift one meter per unit volume of water. Then the mineral processing water costs in a certain hour are expressed as

$$C(\mathbf{Q}) = \sum_{i=1}^n c_i q_i (100 p_i + h_i), \quad (2)$$

where  $h_i$  is the elevation difference of the  $i$ th water source; the integer 100 is used for unit conversion.  $p_i$  is the pressure of the  $i$ th water source. Equation (2) also can be transformed into

$$C(\mathbf{Q}) = \sum_{i=1}^n c_i q_i [100 f(q_{\text{sum}}, q_i, q_i q_j) + h_i], \quad (3)$$

where  $f(q_{\text{sum}}, q_i, q_i q_j)$  is the macro model of the water supply network. This  $q_{\text{sum}}$  is the total flow of mineral processing water required in the mineral processing plant.

The yields of freshwater are low, but its prices are high. So it is important for cost reduction that the recycled water is used in mineral processing. While the total water consumption of mineral processing is met, the recycled water should be used in mineral processing as much as possible. The amount of the recycled water used in mineral processing can be expressed as

$$M(\mathbf{Q}') = M(q_{a+1}, q_{a+2}, \dots, q_n) = \sum_{i=a+1}^b \alpha_i q_i + \sum_{i=b+1}^c \beta_i q_i + \sum_{i=c+1}^n \chi_i q_i, \quad (4)$$

where  $\alpha$  is the penalty factor of tailings water;  $\beta$  is the penalty factor of clean water; and  $\chi$  is the penalty factor of in-plant water.

2.3. *Constraint Condition.* The sum of the flow from the water source should be equal to the total flow of mineral processing water required in mineral processing:

$$\sum_{i=1}^n q_i = q_{\text{sum}}. \quad (5)$$

Each water source should be between the maximum and minimum water supply capacity:

$$q_i^{\min} \leq q_i \leq q_i^{\max} \quad i = 1, 2, \dots, n, \quad (6)$$

where  $q_i^{\max}$  and  $q_i^{\min}$  are upper and lower limits for the flow of the  $i$ th water source.

To ensure efficient transmission, the output pressure of each water source should be restricted as follows:

$$p_i^{\min} \leq p_i \leq p_i^{\max} \quad i = 1, 2, \dots, n, \quad (7)$$

where  $p_i^{\max}$  and  $p_i^{\min}$  are upper and lower limits for the pressure of the  $i$ th water source. Equation (7) also can be transformed into

$$p_i^{\min} \leq f(q_{\text{sum}}, q_i, q_i q_j) \leq p_i^{\max} \quad i = 1, 2, \dots, n. \quad (8)$$

To ensure the pressure requirements of the mineral processing plant, there should be a lower limit for the pressure at the control point.

$$cp_i \geq cp_i^{\min} \quad i = 1, 2, \dots, n, \quad (9)$$

where  $cp_i$  is the pressure at the control point of the  $i$ th water source;  $cp_i^{\min}$  is lower limit for the pressure at the control point of the  $i$ th water source. Equation (9) also can be transformed into

$$f'(q_{\text{sum}}, q_i) \geq cp_i^{\min}, \quad i = 1, 2, \dots, n, \quad (10)$$

where  $f'(q_{\text{sum}}, q_i)$  is the macro model of the pressure control point.

2.4. *Final Model.* In summary, the water optimization model for the mineral processing plant is a two-objective optimization model and its final model is expressed as follows:

$$\begin{aligned} & \min\{C(Q)\}, \max\{M(Q')\}, \\ & \text{s.t.} \begin{cases} \sum_{i=1}^n q_i = q_{\text{sum}}, \\ q_i^{\min} \leq q_i \leq q_i^{\max}, \quad i = 1, 2, \dots, n, \\ p_i^{\min} \leq f(q_{\text{sum}}, q_i, q_i q_j) \leq p_i^{\max}, \quad i = 1, 2, \dots, n, \\ f'(q_{\text{sum}}, q_i) \leq cp_i^{\min}, \quad i = 1, 2, \dots, n. \end{cases} \end{aligned} \quad (11)$$

### 3. Enhancing Multiobjective Artificial Bee Colony Algorithm

With the above model and constraints, the appropriate optimization algorithm is applied to optimize the MPWD problem. In this study, an EMOABC algorithm was proposed to solve the developed MPWD model.

3.1. *Basic ABC Algorithm.* ABC algorithm is a swarm intelligence optimization method proposed by Karaboga to simulate the process of bees searching for nectar [32]. In the framework of the artificial bee colony algorithm, the location of the food source represents the possible candidate solutions of the problem to be optimized, the number of nectars in the food source represents the fitness (objective function value) corresponding to the candidate solutions, and the size of the bee colony is equal to the number of solutions. The basic structure of the ABC algorithm can be divided into the four following stages.

The first is initial stage, where ABC algorithm randomly generates a matrix  $\mathbf{X} = [x_{ij}]_{N \times D}$ ,  $N$  is the population size and represents the number of candidate solutions, and  $D$  is the number of solution parameters,  $i \in (1, 2, \dots, N)$ ,  $j \in (1, 2, \dots, D)$ .  $D$ -dimensional vector  $\mathbf{X}_i = [x_{i1}, x_{i2}, \dots, x_{iD}]$  is one of the candidate solutions.  $x_{ij}$  are randomly generated by

$$x_{ij} = x_j^{\min} + \text{rand}_j(x_j^{\max} - x_j^{\min}), \quad (12)$$

where  $x_j^{\max}$  and  $x_j^{\min}$  are the upper and lower bound constraints of the  $j$ th variable of  $\mathbf{X}_i$ , respectively, and  $\text{rand}_j$  is a uniformly distributed random number within the range [0, 1].

The second is employed bee stage, where an employed bee is to modify and generate a new candidate solution based on the currently collected information, as well as to calculate its fitness. If the fitness of the new candidate solution is higher than the fitness of the original candidate solution, the new candidate solution replaces the original candidate solution; otherwise, the original candidate solution is retained. After all employed bees complete the search task, they share the information on the location of the food source (candidate solution) and the amount of nectar (fitness) they found with onlooker bees. The new candidate solution is generated by

$$v_{ij} = x_{ij} + \varphi_{ij}(x_{ij} - x_{kj}), \quad (13)$$

where  $k$  and  $j$  are randomly chosen indexes,  $k \neq i$ .  $\varphi_{ij}$  is a random number in the range [-1, 1]. The fitness of a solution can be defined as

$$\text{fit}(X_i) = \begin{cases} \frac{1}{1 + f(X_i)}, & \text{if } f(X_i) > 0, \\ 1 + |f(X_i)|, & \text{if } f(X_i) \leq 0. \end{cases} \quad (14)$$

The third is onlooker bee stage, where onlooker bees assess the fitness of candidate solutions passed from the employed bee and then select food sources based on the probability corresponding to the amount of nectar. In this case, the onlooker bee updates the candidate solution based on the original candidate solution and calculates its fitness; if the fitness of the new candidate solution is higher than the original candidate solution, the new candidate solution replaces the original candidate solution; otherwise, the original candidate solution is retained. The probability of a food source chosen by an onlooker bee can be calculated by

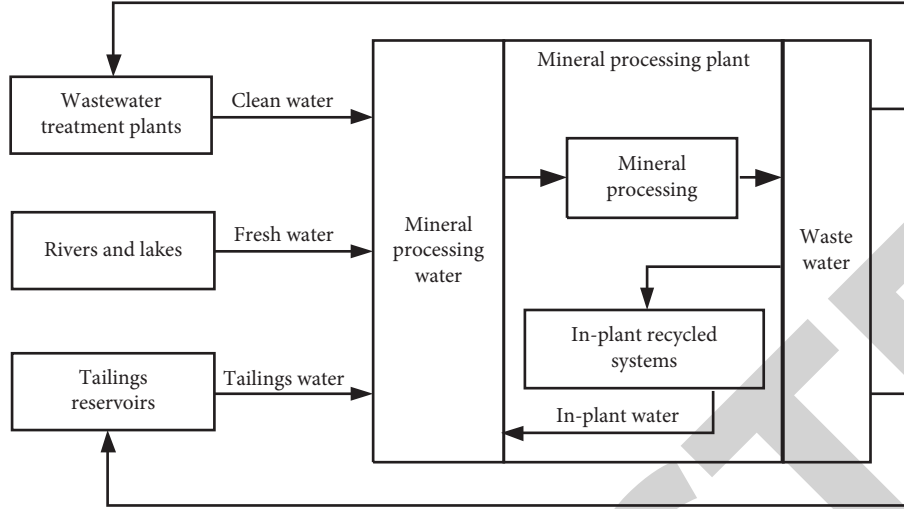


FIGURE 1: The composition of water source in the mineral processing plant.

$$P_i = \frac{\text{fit}(X_i)}{\sum_{i=1}^N \text{fit}(X_i)}. \quad (15)$$

The fourth is scout bee stage; if the nectar of the discovered food source is too inferior to find better solutions several times in a row (“*limit*” is denoted as the maximum number of times), it will be replaced by a new food source that a scout bee randomly found, and then the new food source is generated by (12).

**3.2. Proposed Method.** This section presents a detailed description of the proposed EMOABC algorithm. It can be observed from the literature that this approach to mineral processing water has not yet been extensively explored.

Inspired by the idea of NSGA-II algorithm [30] and DNSPSO algorithm [33], in this paper, the EMOABC is integrated with a series of multiobjective optimization strategies. It includes four important multiobjective strategies: noninferior ranking strategy, Pareto-optimal strategy, crowding distance strategy, and search strategy to evaluate the food source location multiobjectively and select the nondominated solution, so they are presented as follows:

- (1) Noninferior ranking strategy: all individuals in the population that are not dominated by other solutions (noninferior solutions) are defined as rank 1; that is, the virtual fitness of the individual is 1; then these individuals are removed from the population, and new noninferior solutions are identified from the remaining individuals, which are defined as rank 2; that is, the virtual fitness of the individual is 2; the above process is repeated until all individuals in the population have been assigned the corresponding rank. By storing the current solution and all the individuals in the population in a hierarchical manner, the high-performance individuals can have a higher probability of survival and the population rank can be improved rapidly.

- (2) Pareto-optimal strategy: the Pareto set is multiple viable solutions, which can provide multiple viable solutions in the supply of water to the mineral processing plant and can also reduce supply efficiency. Manual Pareto selection suffers from a variety of subjective uncertainties, which can make the selected solution less than optimal.

Define a function  $\eta_i$ , representing the proportion of the  $i$ th objective function value of a solution in the Pareto set that is located along the direction of that objective function in front of the Pareto. The function  $\eta_i$  is expressed as

$$\eta_i = \begin{cases} 1, & V_i \leq V_i^{\min} \\ \frac{V_i^{\max} - V_i}{V_i^{\max} - V_i^{\min}}, & V_i^{\min} \leq V_i \leq V_i^{\max} \\ 0, & V_i \geq V_i^{\max} \end{cases}, \quad (16)$$

where  $V_i^{\max}$  and  $V_i^{\min}$  are the maximum and minimum values of the  $i$ th objective function in the Pareto set, respectively;  $V_i$  is the function value of the  $i$ th objective function. For each noninferior solution  $k$ , the normalized membership function  $\eta^k$  is calculated by the following equation:

$$\eta^k = \frac{\sum_{i=1}^{N_{ob}} \eta_i^k}{\sum_{j=1}^M \sum_{i=1}^{N_{ob}} \eta_i^j}, \quad (17)$$

where  $M$  is the number of noninferior solutions in the Pareto set, and  $N_{ob}$  is the number of objective functions. The larger the value of  $\eta^k$ , the better the performance of  $k$  in coordinating multiple objective functions. Sorting the Pareto set by the  $\eta^k$  value gives a priority sequence of noninferior solutions.

- (3) Crowding distance strategy: the crowding distance is used to estimate how dense a solution is with other solutions around it. For each objective function, the set of noninferior solutions is sorted by the objective function value; then, for each solution  $x$ , the sum of the two sides of the rectangle formed by the solutions  $x+1$  and  $x-1$  is calculated; the final result is the crowding distance  $x_{\text{dist}}$  of the solution  $x$ . The crowding distance of the boundary solution is infinite. When two solutions have the same rank (i.e., the same virtual fitness), the superiority of the two solutions can be compared by the crowding distance. It can be seen that  $x$  individual is superior to  $y$  individual when and only when  $x_{\text{rank}} < y_{\text{rank}}$  or  $x_{\text{rank}} = y_{\text{rank}}$  and  $x_{\text{dist}} > y_{\text{dist}}$ .

- (4) Search strategy: as one of the most representative variants, DNSPSO algorithm is characterized by introducing a global NS operator [33]. Therefore, the search strategy in this paper also adds that operator, and it can be expressed as follows:

$$v_{ij} = r_1 x_{ij} + r_2 x_{\text{best}} + r_3 (x_{aj} - x_{bj}), \quad (18)$$

where  $r_1$ ,  $r_2$ , and  $r_3$  are three mutually exclusive numbers which are randomly chosen from (0, 1), and they have to meet another condition:  $r_1 + r_2 + r_3 = 1$ ;  $x_{\text{best}}$  is the global best solution of the entire population; the indices  $a$  and  $b$  are mutually exclusive integers randomly chosen from (1, 2, ...,  $N$ ), and they are different from the base index  $i$ . It is necessary to point out that  $r_1$ ,  $r_2$ , and  $r_3$  are generated anew in each generation, but they are kept the same for all dimensions in each generation. Once the trial solution is generated, its associated food source has to compete with it for entering the next generation. It implies that the one with a better fitness value has the chance to survive.

Combining the above multiobjective strategies and the ABC algorithm, the basic flow of the EMOABC algorithm is shown in the following:

Step 1: set the number of solution parameters ( $D$ ), the number of populations ( $N$ ), the maximum number of cycles (MNC), and “*limit*” value; randomly generate candidate solutions ( $\mathbf{X} = [x_{ij}]_{N \times D}$ ).

Step 2: calculate the fitness of each objective function, noninferior rank, and the crowding distance.

Step 3: randomly select two solutions and choose the better storage solution. In this way,  $N/2$  better solutions are selected.

Step 4: search in the manner of equation (10) to produce  $N/2$  solution, and if there is a solution beyond the boundary, it is placed on the boundary.

Step 5: identify the better  $N$  solutions from the population of steps 3 and 4 according to noninferior rank and the crowding distance.

Step 6: generate  $N$  new solutions by equation (15), and if there are any solutions beyond the boundary, they are placed on the tenth boundary.

Step 7: select the better  $N$  solutions to form a new population according to noninferior rank and the crowding distance from the solutions in steps 1, 3, and 6.

Step 8: determine if the termination condition is met; if so, output the Pareto front; otherwise, go to step 2.

Step 9: arrange all solutions in the Pareto-optimal set, using equations (13) and (14), and then export a priority list of solutions.

**3.3. Performance of the Proposed Algorithm.** To verify the effectiveness of the EMOABC proposed in this paper, its performance is compared with those of NSGA-II and MOPSO. Convergence and diversity are usually chosen as evaluation criteria, and their specific mathematical expressions can be found in the literature [30, 31]. To ensure the fairness of algorithm testing, the size of the initial population  $N$  is set to 60 and the maximum number of cycles is 1000 for all algorithms on each test function. To achieve the best performance of the other two algorithms, their control parameters were set using the recommended values from the corresponding original literature. For EMOABC, the *limit* was set to  $0.6 \times D \times N$ . To avoid the adverse effect of randomness on algorithm evaluation, each algorithm was run 30 times independently, and the mean value and the standard deviation of convergence and diversity were calculated, and the results of each algorithm for three benchmark functions ZDT1, ZDT2, and ZDT3 are shown in Table 1. Detailed expressions of ZDT1, ZDT2, and ZDT3 can be found in the literature [30]. Three algorithms have been run in MATLAB R2020a on a personal computer with Core 2 Duo 2.13 GHz processor and 4 GB memory.

From Table 1, the best values have been marked in bold. In terms of diversity, compared to other algorithms, the EMOABC algorithm obtained the most uniform Pareto front for each test problem with good distribution characteristics, indicating that the EMOABC algorithm has obvious advantages in distribution characteristics; in terms of convergence, for the ZDT3 problem, the EMOABC algorithm performs slightly worse than the NSGA-II algorithm but better than the MOPSO algorithm. Besides, the standard deviations of convergence and diversity obtained by the EMOABC algorithm are small, indicating that this algorithm is more stable and can almost always converge to the optimal nondominated frontier within the maximum number of cycles, which proves the effectiveness of the EMOABC algorithm to obtain a wider, uniform, and realistic Pareto solution set frontier.

Table 1 cannot completely reflect the results obtained by three algorithms; for this reason, and Figure 2 shows the noninferior optimal solution sets of three benchmark functions obtained by three algorithms.



TABLE 1: The comparable results between EMOABC, NSGA-II, and MOPSO.

Criteria	Algorithm		ZDT1	ZDT2	ZDT3
Convergence	EMOABC	Mean	$3.6239e-3$	$2.2485e-2$	$6.3256e-2$
		Std	$2.3514e-2$	$2.6594e-1$	$2.6585e-3$
	NSGA-II	Mean	$3.9635e-3$	$2.9651e-2$	$5.5883e-2$
		Std	$2.6985e-2$	$2.9985e-1$	$1.4154e-3$
	MOPSO	Mean	$4.1265e-3$	$2.8654e-2$	$7.9562e-2$
		Std	$3.6517e-2$	$3.9914e-1$	$3.1953e-3$
Diversity	EMOABC	Mean	$6.6263e-3$	$5.6156e-2$	$3.6989e-2$
		Std	$4.5696e-4$	$4.2654e-3$	$2.5359e-3$
	NSGA-II	Mean	$3.3659e-2$	$3.4889e-1$	$3.7569e-1$
		Std	$5.6387e-2$	$3.1245e-1$	$3.1245e-2$
	MOPSO	Mean	$6.9159e-2$	$7.6385e-2$	$4.7646e-1$
		Std	$5.4512e-2$	$6.2589e-3$	$3.6632e-2$

As can be seen in Figure 2, the Pareto-optimal solutions of the EMOABC algorithm on ZDT1–3 can approximate the theoretical Pareto-optimal solution of each function very well, which has especially a good distribution. This experimental result once again verifies the effectiveness and advancement of the EMOABC algorithm for the multiobjective optimization problem.

#### 4. Simulation and Discussion

The proposed method has been applied to two cases of mineral processing plants with different water sources and it is compared with MOPSO and NSGA-II to solve this problem. Parameters settings for three algorithms applied to two cases are the same as those set in Section 3.3.

**4.1. Case 1.** Case 1 is that the proposed algorithm is applied to the MPWD model based on mineral processing supplied by four water sources. The parameters about upper and lower limits for Case 1 are shown in Table 2, where 1 represents freshwater, 2 represents tailings water, 3 represents clean water, and 4 represents in-plant water. For recycled water, the lower cost with water source should be used first. In Case 1, since  $c_2 < c_4 < c_3$ , set the penalty factor  $\alpha = 3$ ,  $\beta = 14$ , and  $\chi = 8$ . The macro model, which uses historical data as a sample, is built by a stepwise regression approach [34]. The macro model of the water supply network and the macro model of the pressure control point in Case 1 are given in Tables 3 and 4, respectively. The following macro models (in Tables 3–6) are all simulation models.

To optimize both mineral processing water costs and the amount of recycled water simultaneously, the proposed EMOABC algorithm has been applied in the MPWD model with  $q_{\text{sum}} = 3000$  and  $4000$ . The five best solutions and best  $\eta k$  are summarized in Tables 7 and 8. The higher the value of  $\eta k$  for a solution, the higher the priority of the solution. From Table 7, we can see that  $C(Q)$  in  $S_2$  and  $S_3$  is slightly different, but  $q_1$ – $q_4$  in  $S_2$  and  $S_3$  are completely different. Also, the difference between  $q_1$ – $q_4$  in  $S_2$  and  $S_5$  is not significant, but the difference of  $C(Q)$  in  $S_2$  and  $S_5$  is obvious. The above denotes that a set of optimal solutions are obtained by the proposed EMOABC algorithm. The comparison between Tables 7 and 8 shows that when  $q_{\text{sum}}$  rises

from 3000 to 4000,  $q_1$ ,  $q_2$ ,  $q_3$ , and  $q_4$  increase accordingly, while  $q_3$  generally declines.  $q_2$  increases the most due to its low cost, and the decline in  $q_3$  is due to its high cost. The results of using EMOABC algorithm to schedule the flow of different water sources are consistent with those of manual regulation.

The MPWD model is solved using EMOABC, NSGA-II, and MOPSO. A comparison of the results obtained by three algorithms is shown in Table 9. It is seen that the EMOABC algorithm yields a lower cost of mineral processing water and a higher amount of recycled water (in bold) than the other two algorithms.

From Figure 3, it can be observed that EMOABC can discover a well-distributed and diverse solution set for this problem, while the other two algorithms only find some sparse distributions.

**4.2. Case 2.** Case 2 is that the proposed algorithm is applied to the MPWD model based on mineral processing supplied by six water sources. The parameters about upper and lower limits for Case 2 are shown in Table 10, where 1 represents freshwater no. 1, 2 represents freshwater no. 2, 3 represents tailings water no. 1, 4 represents tailings water no. 2, 5 represents clean water, and 6 represents in-plant water. For recycled water, the water source with a lower cost should be used first. In Case 2, since  $c_3 < c_4 < c_6 < c_5$ , set the penalty factor  $\alpha_1 = 3$ ,  $\alpha_2 = 4$ ,  $\beta = 14$ , and  $\chi = 8$ . The macro model of the water supply network and the macro model of the pressure control point in Case 2 are given in Tables 5 and 6.

In this case,  $C(Q)$  and  $M(Q')$ , two competing objectives, are optimized simultaneously by the EMOABC algorithm, and the five best solutions are shown in Tables 11 and 12. From Tables 11 and 12, the convergence of the solutions in Case 2 is as good as that in Case 1. It is worth noting that as the number of water sources increases, the water selection options are diverse. A comparison of the results obtained by MOPSO, NSGA-II, and EMOABC in Case 2 is shown in Table 13. From Table 13, we can observe that the EMOABC algorithm obtains the optimal solution (in bold) in Case 2 similar to Case 1.

Figure 4 shows the distribution of nondominated solutions obtained by NSGA-II, MOPSO, and EMOABC in Case 2. It is seen from them that EMOABC spreads well

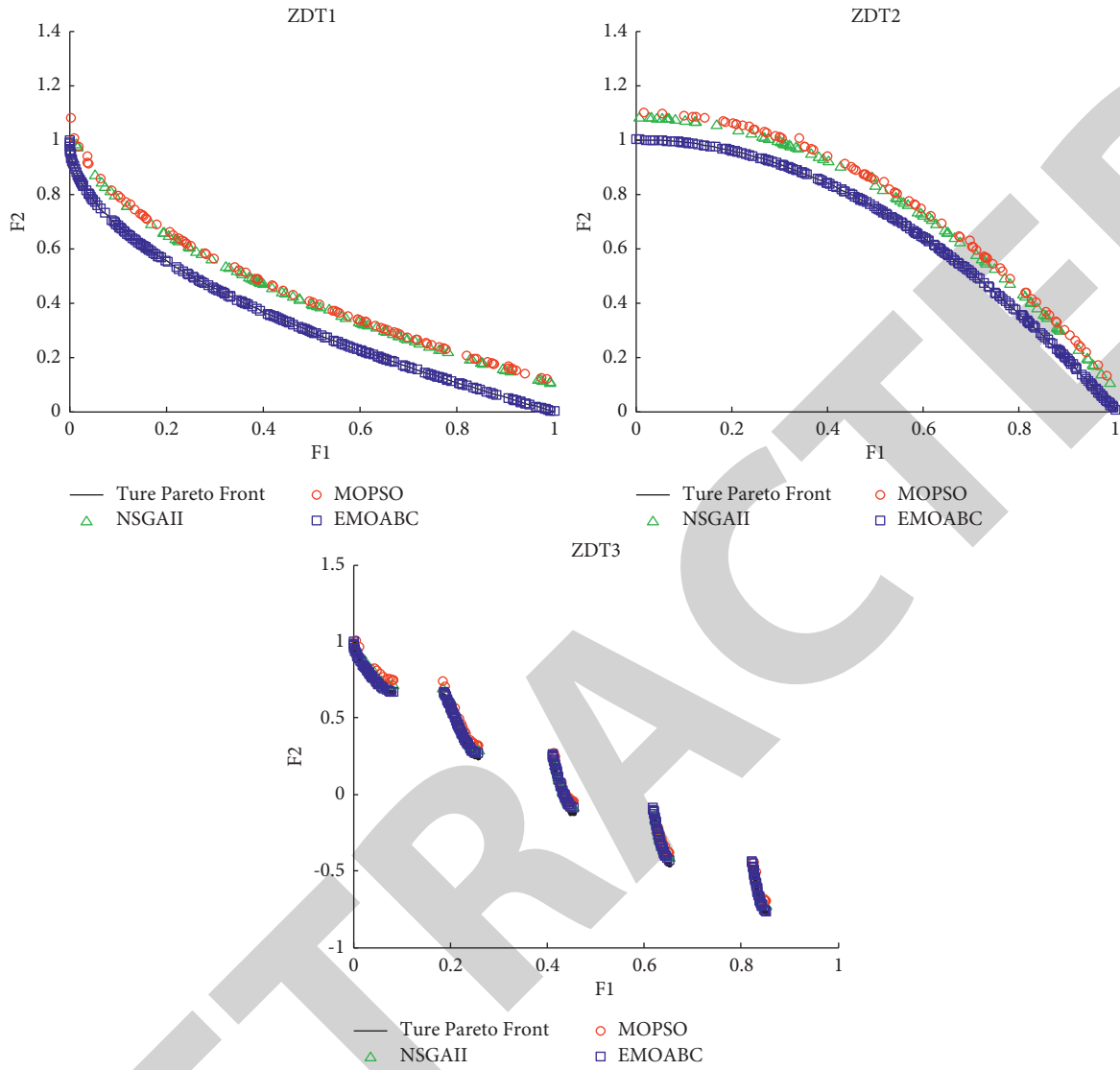


FIGURE 2: The comparable Pareto front obtained by three algorithms.

TABLE 2: Related parameters in Case 1.

Water sources	$q_i^{\min}$ (m <sup>3</sup> /h)	$q_i^{\max}$ (m <sup>3</sup> /h)	$p_i^{\min}$ (MPa)	$p_i^{\max}$ (MPa)	$h_i$ (m)	$c_i$ (\$/m <sup>3</sup> )	$cp_i^{\min}$ (MPa)
1	100	1000	0.15	0.95	15	0.11	0.13
2	120	3000	0.18	0.86	56	0.24	0.14
3	110	2000	0.17	0.93	26	0.46	0.13
4	95	1500	0.16	0.74	5	0.32	0.11

TABLE 3: Related macro model of water supply network in Case 1.

Water sources	$f(q_{\text{sum}}, q_i, q_j q_j)$
1	$0.145 + 2.618 \times 10^{-9} \times q_1^2 + 3.698 \times 10^{-9} \times q_3^2 + 6.594 \times 10^{-8} \times q_2 q_3 + 2.365 \times 10^{-8} \times q_{\text{sum}}^2$
2	$0.128 + 4.513 \times 10^{-9} \times q_1^2 + 6.485 \times 10^{-8} \times q_2^2 + 2.684 \times 10^{-10} \times q_1 q_2 + 9.264 \times 10^{-10} \times q_{\text{sum}}^2$
3	$0.117 + 1.561 \times 10^{-9} \times q_3^2 + 1.248 \times 10^{-9} \times q_4^2 + 5.654 \times 10^{-9} \times q_1 q_4 + 4.589 \times 10^{-8} \times q_{\text{sum}}^2$
4	$0.124 + 2.415 \times 10^{-9} \times q_4^2 + 8.261 \times 10^{-9} \times q_1 q_3 + 1.368 \times 10^{-8} \times q_{\text{sum}}^2$



TABLE 4: Related macro model of the pressure control point in Case 1.

Water source	$f'(q_{\text{sum}}, q_i)$
1	$0.109 + 1.529 \times 10^{-9} \times q_1^2 + 3.659 \times 10^{-10} \times q_{\text{sum}}^2$
2	$0.123 + 4.268 \times 10^{-8} \times q_2^2 + 5.364 \times 10^{-9} \times q_{\text{sum}}^2$
3	$0.151 + 7.268 \times 10^{-10} \times q_3^2 + 5.269 \times 10^{-8} \times q_{\text{sum}}^2$
4	$0.124 + 6.324 \times 10^{-8} \times q_4^2 + 8.624 \times 10^{-10} \times q_{\text{sum}}^2$

TABLE 5: Related macro model of water supply network in Case 2.

Water sources	$f(q_{\text{sum}}, q_i, q_i q_j)$
1	$0.125 + 3.584 \times 10^{-8} \times q_1^2 + 5.629 \times 10^{-10} \times q_2^2 + 2.598 \times 10^{-9} \times q_3 q_4 + 4.268 \times 10^{-9} \times q_{\text{sum}}^2$
2	$0.114 + 8.268 \times 10^{-10} \times q_2^2 + 3.628 \times 10^{-9} \times q_6^2 + 6.324 \times 10^{-9} \times q_3 q_5 + 8.529 \times 10^{-10} \times q_{\text{sum}}^2$
3	$0.115 + 1.269 \times 10^{-8} \times q_3^2 + 1.562 \times 10^{-8} \times q_6^2 + 4.263 \times 10^{-9} \times q_4 q_5 + 9.264 \times 10^{-10} \times q_{\text{sum}}^2$
4	$0.142 + 7.261 \times 10^{-9} \times q_4^2 + 6.826 \times 10^{-9} \times q_1^2 + 5.987 \times 10^{-9} \times q_1 q_3 + 8.154 \times 10^{-10} \times q_{\text{sum}}^2$
5	$0.138 + 9.261 \times 10^{-10} \times q_5^2 + 5.698 \times 10^{-10} \times q_2^2 + 4.358 \times 10^{-9} \times q_3 q_6 + 7.569 \times 10^{-9} \times q_{\text{sum}}^2$
6	$0.152 + 3.652 \times 10^{-8} \times q_6^2 + 1.268 \times 10^{-8} \times q_2 q_4 + 2.584 \times 10^{-9} \times q_{\text{sum}}^2$

TABLE 6: Related macro model of the pressure control point in Case 2.

Water source	$f'(q_{\text{sum}}, q_i)$
1	$0.121 + 2.316 \times 10^{-10} \times q_1^2 + 8.362 \times 10^{-9} \times q_{\text{sum}}^2$
2	$0.132 + 5.326 \times 10^{-9} \times q_2^2 + 3.958 \times 10^{-10} \times q_{\text{sum}}^2$
3	$0.142 + 9.635 \times 10^{-10} \times q_3^2 + 4.369 \times 10^{-9} \times q_{\text{sum}}^2$
4	$0.129 + 2.589 \times 10^{-8} \times q_4^2 + 5.241 \times 10^{-10} \times q_{\text{sum}}^2$
5	$0.161 + 7.125 \times 10^{-9} \times q_5^2 + 8.629 \times 10^{-10} \times q_{\text{sum}}^2$
6	$0.135 + 5.264 \times 10^{-9} \times q_6^2 + 1.259 \times 10^{-8} \times q_{\text{sum}}^2$

TABLE 7: Five best solutions for Case 1 using the EMOABC ( $q_{\text{sum}} = 3000$ ).

Solutions	$q_1$ (m <sup>3</sup> /h)	$q_2$ (m <sup>3</sup> /h)	$q_3$ (m <sup>3</sup> /h)	$q_4$ (m <sup>3</sup> /h)	$C$ (Q) (\$/h)	$M$ (Q') (m <sup>3</sup> /h)	$\eta k$
$S_1$	328	1156	875	641	61577	20084	0.0536
$S_2$	396	1297	801	506	61683	19153	0.0534
$S_3$	298	1307	783	612	61681	19779	0.0531
$S_4$	384	1376	752	488	61613	18560	0.0529
$S_5$	395	1298	803	504	61756	19168	0.0528

TABLE 8: Five best solutions for Case 1 using the EMOABC ( $q_{\text{sum}} = 4000$ ).

Solutions	$q_1$ (m <sup>3</sup> /h)	$q_2$ (m <sup>3</sup> /h)	$q_3$ (m <sup>3</sup> /h)	$q_4$ (m <sup>3</sup> /h)	$C$ (Q) (\$/h)	$M$ (Q') (m <sup>3</sup> /h)	$\eta k$
$S_1$	422	2120	804	654	103820	22848	0.0546
$S_2$	439	2206	739	616	103450	21892	0.0554
$S_3$	432	2072	824	672	103380	23128	0.0562
$S_4$	398	2218	730	654	103580	22106	0.0558
$S_5$	406	2195	752	647	103790	22289	0.0549

TABLE 9: The best solutions for Case 1 using three algorithms.

	MOPSO		NSGA-II		EMOABC	
	$q_{\text{sum}} = 3000$	$q_{\text{sum}} = 4000$	$q_{\text{sum}} = 3000$	$q_{\text{sum}} = 4000$	$q_{\text{sum}} = 3000$	$q_{\text{sum}} = 4000$
$q_1$	395	419	329	395	384	432
$q_2$	1294	2118	1159	2221	1376	2082
$q_3$	802	807	864	732	752	794
$q_4$	509	656	648	652	488	692
$C$ (Q)	61668	103900	61641	103750	<b>61577</b>	<b>103380</b>
$M$ (Q')	19182	22900	20057	22127	<b>20084</b>	<b>23128</b>

Bold highlights that the algorithm proposed in this paper achieves better results.

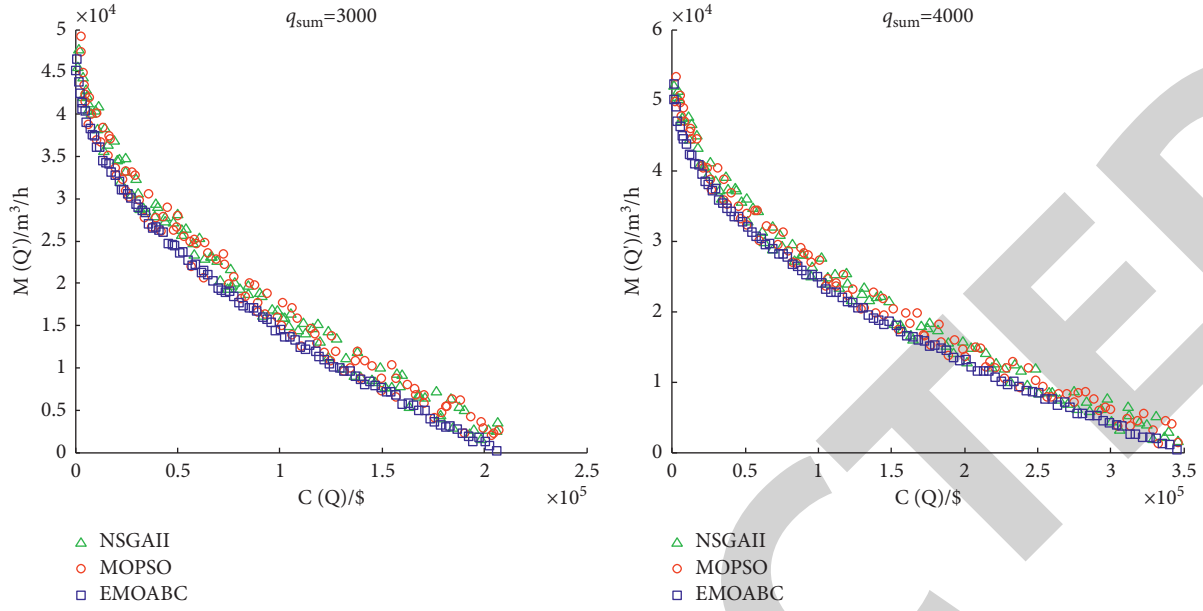


FIGURE 3: Pareto fronts obtained by NSGA-II, MOPSO, and EMOABC for  $C(Q)$  and  $M(Q')$  in Case 1.

TABLE 10: Related parameters in Case 2.

Water sources	$q_i^{\min}$ (m <sup>3</sup> /h)	$q_i^{\max}$ (m <sup>3</sup> /h)	$p_i^{\min}$ (MPa)	$p_i^{\max}$ (MPa)	$h_i$ (m)	$c_i$ (\$/m <sup>3</sup> )	$cp_i^{\min}$ (MPa)
1	110	900	0.16	0.96	12	0.12	0.14
2	100	800	0.19	0.87	23	0.13	0.15
3	120	2500	0.18	0.92	45	0.24	0.16
4	105	2800	0.17	0.76	56	0.26	0.13
5	130	1200	0.15	0.85	25	0.48	0.17
6	115	950	0.14	0.63	3	0.33	0.14

TABLE 11: Five best solutions for Case 2 using EMOABC ( $q_{\text{sum}} = 5000$ ).

Solutions	$q_1$ (m <sup>3</sup> /h)	$q_2$ (m <sup>3</sup> /h)	$q_3$ (m <sup>3</sup> /h)	$q_4$ (m <sup>3</sup> /h)	$q_5$ (m <sup>3</sup> /h)	$q_6$ (m <sup>3</sup> /h)	$C(Q)$ (\$/h)	$M(Q')$ (m <sup>3</sup> /h)	$\eta k$
$S_1$	218	324	1487	1295	894	692	80912	27693	0.0585
$S_2$	352	415	1352	1415	841	625	80005	26490	0.0587
$S_3$	364	325	1238	1449	828	814	80087	27614	0.0596
$S_4$	316	348	1336	1479	779	742	80049	26766	0.0594
$S_5$	355	362	1302	1209	893	879	79153	28276	0.0598

TABLE 12: Five best solutions for Case 2 using EMOABC ( $q_{\text{sum}} = 6000$ ).

Solutions	$q_1$ (m <sup>3</sup> /h)	$q_2$ (m <sup>3</sup> /h)	$q_3$ (m <sup>3</sup> /h)	$q_4$ (m <sup>3</sup> /h)	$q_5$ (m <sup>3</sup> /h)	$q_6$ (m <sup>3</sup> /h)	$C(Q)$ (\$/h)	$M(Q')$ (m <sup>3</sup> /h)	$\eta k$
$S_1$	402	406	1789	1703	906	794	102800	31215	0.0642
$S_2$	421	456	1859	1906	806	552	102540	28901	0.0649
$S_3$	359	396	1952	1589	863	841	102300	31022	0.0651
$S_4$	398	412	1652	1856	786	896	102240	31272	0.0648
$S_5$	416	395	1742	1647	911	889	102230	31680	0.0654

on the optimization problem, the function value obtained by EMOABC is significantly better than that obtained by NSGA-II, and MOPSO and EMOABC can find a better spread and more solutions in the entire Pareto-optimal region than the other two algorithms. It is indicated that the proposed algorithm is more effective than the other

algorithms for solving the MPWD problem. Figures 3 and 4 also show that the two objectives of maximizing the amount of recycled water and minimizing mineral processing water costs conflict. As mineral processing water costs increase, the recycled water consumption decreases, and vice versa.

TABLE 13: The best solutions for Case 2 using three algorithms.

	MOPSO		NSGA-II		EMOABC	
	$q_{\text{sum}} = 5000$	$q_{\text{sum}} = 6000$	$q_{\text{sum}} = 5000$	$q_{\text{sum}} = 6000$	$q_{\text{sum}} = 5000$	$q_{\text{sum}} = 6000$
$q_1$	322	415	388	429	355	416
$q_2$	405	323	361	398	362	395
$q_3$	1212	1842	1251	1662	1302	1742
$q_4$	1315	1772	1289	1731	1209	1647
$q_5$	881	845	896	919	893	911
$q_6$	865	803	815	861	879	889
$C(Q)$	79379	102960	79398	102550	<b>79153</b>	<b>102230</b>
$M(Q')$	28150	30868	27973	31664	<b>28276</b>	<b>31680</b>

Bold highlights that the algorithm proposed in this paper achieves better results.

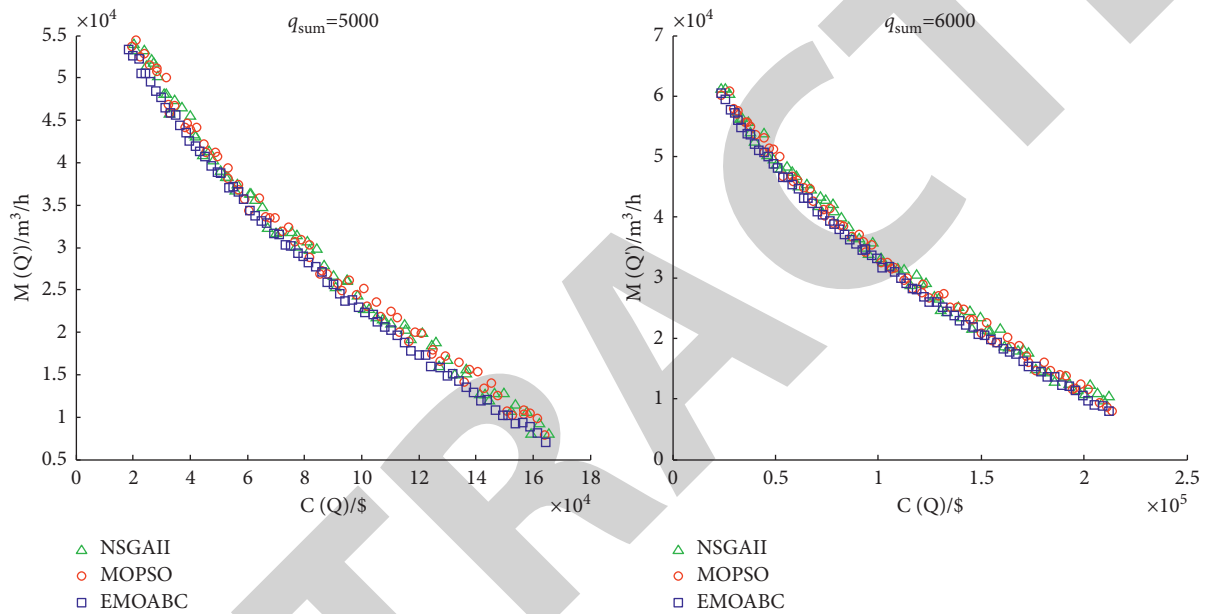


FIGURE 4: Pareto fronts obtained by NSGA-II, MOPSO, and EMOABC for  $C(Q)$  and  $M(Q')$  in Case 2.

TABLE 14: Computing time of all algorithms for both cases.

	Computing time (s)		
	EMOABC	NSGA-II	MOPSO
Case 1	58.6	40.2	47.8
Case 2	74.4	59.5	64.7

**4.3. Algorithm Complexity Analysis.** Complexity analysis of the above algorithms was performed using the time response for both cases. The average computation time for all algorithms over 30 runs is given in Table 14. From Table 14, it can be seen that EMOABC takes the most computation time. This can be explained by the fact that the multiple strategies integrated in EMOABC enhance the search capability at the cost of increased computation. In conclusion, EMOABC requires more computation time to achieve better results compared to other algorithms.

## 5. Conclusions

The main goal of this study was to solve the MPWD problem effectively. In this study, an MPWD optimization model was developed to maximize the amount of recycled water and minimize mineral processing water costs, with consideration of various constraints. To solve this multiobjective model, the EMOABC algorithm was proposed as an optimizer derived from the ABC algorithm. This study has shown that the developed model was theoretically valid and reasonable

and could be utilized to provide decision support for actual MPWD. The study has also shown that the proposed algorithm is competent to solve the MPWD problem and is better than NSGA-II and MOPSO in both benchmark functions and cases.

A limitation of this study is that the developed model only has two objectives. The MPWD problem described as a three-objective or even four-objective model may be more realistic. Therefore, building a three-objective or even four-objective model for this MPWD problem is a matter of future work. Meanwhile, continued research of the multiobjective artificial bee colony (MOABC) algorithm is also a focus. In particular, aiming at a three-objective or even four-objective model for this MPWD problem, the concept of multihives bee will be introduced into the ABC algorithm to further improve the performance of the MOABC algorithm.

### Data Availability

No data were used to support this study.

### Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

### Acknowledgments

This project was fully supported by Liaoning Provincial Department of Education Basic Research Projects for Higher Education Institutions, China (no. LJKZ0301), the Scientific Research Foundation of the Education Department of Liaoning Province, China (no. 2017LNQN22), and Young Teachers Foundation of University of Science and Technology Liaoning, China (no. 2017QN04).

### References

- [1] P. R. Bhave and V. V. Sonak, "A critical study of the linear programming gradient method for optimal design of water supply networks," *Water Resources Research*, vol. 28, no. 6, pp. 1577–1584, 1992.
- [2] J. A. Frizzone, R. D. Coelho, D. Dourado-Neto, and R. Soliani, "Linear programming model to optimize the water resource use in irrigation projects: an application to the Senator Nilo Coelho Project," *Scientia Agricola*, vol. 54, no. spe, pp. 136–148, 1997.
- [3] G. Yu, R. S. Powell, and M. J. H. Sterling, "Optimized pump scheduling in water distribution systems," *Journal of Optimization Theory and Applications*, vol. 83, no. 3, pp. 463–488, 1994.
- [4] H. M. V. Samani and A. Mottaghi, "Optimization of water distribution networks using integer linear programming," *Journal of Hydraulic Engineering*, vol. 132, no. 5, pp. 501–509, 2006.
- [5] A. Haghghi, H. M. V. Samani, and Z. M. V. Samani, "GA-ILP method for optimization of water distribution networks," *Water Resources Management*, vol. 25, no. 7, pp. 1791–1808, 2011.
- [6] C. Zoppou and M. Reed, "Water distribution network design optimization: simulated annealing approach," *Journal of Water Resources Planning and Management*, vol. 127, no. 1, pp. 69–70, 2001.
- [7] J. Marques, M. Cunha, and D. A. Savić, "Multi-objective optimization of water distribution systems based on a real options approach," *Environmental Modelling & Software*, vol. 63, pp. 1–13, 2015.
- [8] K. Zhang, H. Yan, H. Zeng, K. Xin, and T. Tao, "A practical multi-objective optimization sectorization method for water distribution network," *The Science of the Total Environment*, vol. 656, pp. 1401–1412, 2019.
- [9] B. A. Tolson, H. R. Maier, A. R. Simpson, and B. J. Lence, "Genetic algorithms for reliability-based optimization of water distribution systems," *Journal of Water Resources Planning and Management*, vol. 130, no. 1, pp. 63–72, 2004.
- [10] N. Moosavian and B. J. Lence, "Nondominated sorting differential evolution algorithms for multiobjective optimization of water distribution systems," *Journal of Water Resources Planning and Management*, vol. 143, no. 4, Article ID 04016082, 2017.
- [11] R. Mansouri, H. Torabi, M. Hoseini, and H. Morshedzadeh, "Optimization of the water distribution networks with differential evolution (DE) and mixed integer linear programming (MILP)," *Journal of Water Resource and Protection*, vol. 7, no. 9, pp. 715–729, 2015.
- [12] N. Moosavian and B. J. Lence, "Fittest individual referenced differential evolution algorithms for optimization of water distribution networks," *Journal of Computing in Civil Engineering*, vol. 33, no. 6, Article ID 04019036, 2019.
- [13] S. Pandey and Y. P. Mathur, "Optimization of water distribution network using particle swarm optimization," *International Journal of Applied Engineering Research*, vol. 8, no. 12, pp. 92–101, 2013.
- [14] I. Montalvo, J. Izquierdo, S. Schwarze, and R. Perez Garcia, "Multi-objective particle swarm optimization applied to water distribution systems design: an approach with human interaction," *Mathematical and Computer Modelling*, vol. 52, no. 7–8, pp. 1219–1227, 2010.
- [15] X. Qi, K. Li, and W. D. Potter, "Estimation of distribution algorithm enhanced particle swarm optimization for water distribution network optimization," *Frontiers of Environmental Science & Engineering*, vol. 10, no. 2, pp. 341–351, 2016.
- [16] A. C. Zecchin, H. R. Maier, A. R. Simpson, M. Leonard, and J. B. Nixon, "Ant colony optimization applied to water distribution system design: comparative study of five algorithms," *Journal of Water Resources Planning and Management*, vol. 133, no. 1, pp. 87–92, 2007.
- [17] A. Afshar, F. Massoumi, A. Afshar, and M. A. Mariño, "State of the art review of ant colony optimization applications in water resource management," *Water Resources Management*, vol. 29, no. 11, pp. 3891–3904, 2015.
- [18] M. López-Ibáñez, T. D. Prasad, and B. Paechter, "Ant colony optimization for optimal control of pumps in water distribution networks," *Water Resour Plan Manag*, vol. 134, no. 4, pp. 337–346, 2008.
- [19] W. Kurek and A. Ostfeld, "Multi-objective optimization of water quality, pumps operation, and storage sizing of water distribution systems," *Journal of Environmental Management*, vol. 115, pp. 189–197, 2013.
- [20] M. M. Eusuff and K. E. Lansey, "Optimization of water distribution network design using the shuffled frog leaping algorithm," *Journal of Water Resources Planning and Management*, vol. 129, no. 3, pp. 210–225, 2003.
- [21] D. Mora-Melia, P. Iglesias-Rey, F. Martínez-Solano, and P. Muñoz-Velasco, "The efficiency of setting parameters in a

## Retraction

# Retracted: Novel Decision-Making Techniques in Tripolar Fuzzy Environment with Application: A Case Study of ERP Systems

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.




The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] M. Afridi, A. H. Gumaei, H. AlSalman, A. Khan, and S. M. Mizanur Rahman, "Novel Decision-Making Techniques in Tripolar Fuzzy Environment with Application: A Case Study of ERP Systems," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4488576, 23 pages, 2022.

## Research Article

# Novel Decision-Making Techniques in Tripolar Fuzzy Environment with Application: A Case Study of ERP Systems

Minhaj Afridi,<sup>1</sup> Abdu H. Gumaei ,<sup>2</sup> Hussain ALSalman ,<sup>3</sup> Asghar Khan ,<sup>1</sup> and Sk. Md. Mizanur Rahman<sup>4</sup>

<sup>1</sup>Department of Mathematics, Faculty of Physical and Numerical Sciences, Abdul Wali Khan University, Mardan, Khyber Pakhtunkhwa, Pakistan

<sup>2</sup>Department of Computer Science, Faculty of Applied Sciences, Taiz University, Taiz 6803, Yemen

<sup>3</sup>Department of Computer Science, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia

<sup>4</sup>Information and Communication Engineering Technology, School of Engineering Technology and Applied Science, Centennial College, Toronto, Canada

Correspondence should be addressed to Abdu H. Gumaei; [abdugumaei@taiz.edu.ye](mailto:abdugumaei@taiz.edu.ye)

Received 1 October 2021; Revised 21 December 2021; Accepted 22 December 2021; Published 31 January 2022

Academic Editor: Daqing Gong

Copyright © 2022 Minhaj Afridi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The intuitionistic fuzzy set (IFS) and bipolar fuzzy set (BFS) are all effective models to describe ambiguous and incomplete cognitive knowledge with membership, non-membership, negative membership, and hesitancy sections. But in daily life problems, there are some situations where we cannot apply the ordinary models of IFS and BFS, separately. Hence, there is a need to combine both the models of IFS and BFS into a single one. A tripolar fuzzy set (TFS) is a generalization of IFS and BFS. In circumstances where BFS and IFS models cannot be used individually, a tripolar fuzzy model is more dependable and efficient. Further, the IFS and BFS models are reduced to corollaries due to the proposed model of TFS. For this purpose in this article, we first consider some novel operations on tripolar fuzzy information. These operations are formulated on the basis of well-known Dombi T-norm and T-conorm, and the desirable properties are discussed. By applying the Dombi operations, arithmetic and geometric aggregation operators of TFS are proposed, and we introduce the concepts of a TF-Dombi weighted average (TFDWA) operator, a TF-Dombi ordered weighted average (TFDOWA) operator, and a TF-Dombi hybrid weighted (TFDHW) operator and explore their fundamental features including idempotency, boundedness, monotonicity, and others. In the second part, we propose TF-Dombi weighted geometric (TFDWG) operator, TF-Dombi ordered weighted geometric (TFDOWG) operator, and TF-Dombi hybrid geometric (TFDHG) operator. The features and specific cases of the mentioned operators are examined. Enterprise resource planning (ERP) is a management and integration approach that organizations employ to manage and develop many aspects of their operations. The study's primary contribution is to employ TFS to create certain decision-making strategies for the selection of optimal ERP systems. The proposed operators are then used to build several techniques for solving multiattribute decision-making (MADM) issues with TF information. Finally, an example of ERP system selection is investigated to demonstrate that the techniques suggested are trustworthy and realistic.

## 1. Introduction

Decision making (DM) is a method of resolving real-world problems by selecting the ideal choice from a range of viable options. MADM is an area of operations research in which the best answer is found after weighing all of the options against a set of criteria. In real life, there are numerous

challenges that are ambiguous and uncertain. To deal with uncertain and ambiguous information, Zadeh developed fuzzy set (FS) theory [1]. The idea of intuitionistic fuzzy set (IFS) developed by Atanassov [2, 3] comprised two degree membership functions, namely, accepting and denying. IFS is a potential generalization of the notion of a FS [1], whose degree component contained only accepting degree.



Decision science is a continuously growing ground in the modern technological era. In the decision process, experts select an optimal alternative under some preference values imposed by the experts in a given finite set of alternatives. Decision-making methods have been used in several areas of modern science, for example, Xu [4] developed the use of IFS in arithmetic aggregation operators (AOPs) and initiated many valuable operators. Afterwards, Xu and Yager [5] developed some geometric AOPs and demonstrated the use of these operators in MADM. Since its inception, the IFS has attracted a lot of interest, including dynamic MADM in the IF setting [6, 7], IF aggregation operators [8–13], IF entropy [14–16], IF generalized Dice similarity measures [16], IF TOPSIS [14–16], and IF gray relational analysis [17–19]. Many generalizations of IFS were developed in the literature after the successful stages of IFS. The bipolar fuzzy set (BFS) [20, 21] was developed to measure the uncertain and cognitive information presented in the form of positive polarity and negative polarity in real-world scenarios. The objects in a universe in BFS are characterized by positive polarity and negative polarity, and the BFS has range of membership in  $[-1, 1]$ . Chen et al. [22] studied an empirical examination of attribute interrelationships that are heterogeneous in a MADM. For MADM, a unique consensus model based on multigranular HFLTSs was established in [23]. In [24], a method was developed and implemented in MADM issues to present a linguistic distribution assessment (LDA) using a hesitant linguistic distribution (HLD). A new strategy for dealing with MCGDM problems with unbalanced HFLTSs was established in [25] taking psychological behaviour of DMs. The notion of a BFS has been employed in many potential areas including bipolar logical reasoning and other set theoretical abstract structures [20], theoretical approach to traditional medicine of China [26], computational psychiatry [27], decision analysis and organizational modeling [28], and quantum computing [29]. Recently, Dombi [30] studied decision methods by developing some arithmetic and geometric AOPs with the help of Dombi operations and BFSs. The topics in fuzzy information aggregation operators are developing rapidly, and many researchers are involved to construct feasible and advanced models to deal with decision processes. In [31], Liu applied Hamacher AOPs in interval-valued IF numbers (IVIFNs) and constructed MAGDM methods. Wang and Garg [32] defined some Pythagorean fuzzy interaction aggregation operators with the aid of Archimedean t-conorm and t-norm (ATT) to aggregate the numbers. Zhang [33] proposed IVIFNs in Frank AOPs and considered the applications in MAGDM. Zhang and Zhang [34] defined Einstein hybrid AOPs for IFNs and applied it to the MADM method.

A new fuzzy extension has just been developed, dubbed tripolar FS (TFS), which is a generalization of IFS and BFS [35]. In instances where IFS and BFS models are difficult to apply, a TF model can be used. In TFS, we consider a triplet of real numbers, namely, the membership, non-membership, and negative membership degrees, which is used to define an object in a TFS. Similar to BFS, the range of membership of the TFS is also  $[-1, 1]$ . A TFS model can easily be applied in situations where IFS and BFS models fail.

Wei [8] developed several novel operations known as Dombi T-norm (DTN) and Dombi T-conorm (DTCN), which have a high potential for parameter variation. Han et al. [26] took advantage by performing Dombi operations on IFSs and developing MAGDM problems utilizing the Dombi Bonferroni mean operator and IF information. Wei et al. [21, 36], in a single-valued neutrosophic environment, constructed a MADM problem utilizing Dombi operations. Lu and Busemeyer [27] defined typhoon disaster assessment using Dombi operations in a HF context. In this paper, we presented various AOPs in a TF environment by using an expanded idea of IFS and BFS. The following points describe the novelty of proposed operators:

- (i) The ability of a TFS is to express IFS and BFS information at once in a single notion called tripolar fuzzy environment which makes it exceptional in literature. The qualitative characteristics of IFS and BFS are combined in a single TFS. As a result, the work is presented in a TFS context to deal with tripolarity type of fuzzy information.
- (ii) The flexibility parameter involved in Dombi operations has the ability to produce more accurate results in a decision process.
- (iii) The proposed model can be applied in situations where the traditional models of IFS and BFS fail.
- (iv) Dombi operations' flexible parameter make it simple to investigate the stability of ranking order of alternatives.
- (v) The score function was crucial in creating a ranking order among the choices in DM situations. By integrating the concepts of IFS and BFS score functions, a new concept of score function is constructed in the proposed study.

The rest of the paper is structured as follows.

The essential ideas of the IFS, BFSs, and TFSs, as well as the operational regulations of TFNs, will be addressed in the next section. In Section 3, we established the TF-Dombi weighted average (TFDWA), TF-Dombi ordered weighted average (TFDOWA), TF-Dombi hybrid average (TFDHA), TF-Dombi weighted geometric (TFDWG), TF-Dombi ordered weighted geometric (TFDOWG), and TF-Dombi hybrid geometric (TFDHG) operators. The peculiarities and specific cases of these operators are also explored. In Section 4, we used these operators to come up with some solutions to the TF MADM challenges. Section 5 examines an illustrated example of ERP system selection. Some remarks are made in Section 6, where we conclude the article.

## 2. Preliminaries

In this section, we review the definitions of IFS, BFS, and TFS extracted from [1–3, 37], and then we construct a novel notion of score and accuracy functions and introduce a new comparison technique for TFS.

*Definition 1* (see [2, 3]). An IFS  $f$  of a non-empty set  $\tilde{X}$  is an object of the form

$$\tilde{f} = \langle \mu_{\tilde{f}}^-, \lambda_{\tilde{f}}^- \rangle = \left\{ \langle x, \mu_{\tilde{f}}^-(x), \lambda_{\tilde{f}}^-(x) \rangle \mid x \in \tilde{X} \right\}, \quad (1)$$

where  $\mu_{\tilde{f}}^-: \tilde{X} \rightarrow [0, 1]$  and  $\lambda_{\tilde{f}}^-: \tilde{X} \rightarrow [0, 1]$  represent the degree of membership and non-membership of an element  $x$  in FS  $\tilde{f}$ , respectively, and  $0 \leq \mu_{\tilde{f}}^-(x) + \lambda_{\tilde{f}}^-(x) \leq 1$  for all  $x \in \tilde{X}$ .

**Definition 2** (see [37]). A BFS  $A$  of a non-empty set  $\tilde{X}$  is an object with a shape

$$\tilde{A} = \langle \mu_{\tilde{A}}^-, \delta_{\tilde{A}}^- \rangle = \left\{ \langle x, \mu_{\tilde{A}}^-(x), \delta_{\tilde{A}}^-(x) \rangle \mid x \in \tilde{X} \right\}, \quad (2)$$

where  $\mu_{\tilde{A}}^-: \tilde{X} \rightarrow [0, 1]$  and  $\delta_{\tilde{A}}^-: \tilde{X} \rightarrow [-1, 0]$  indicate the degree to which an element  $x$  satisfies the associated property to BFS  $\tilde{A}$ , as well as the implicit counter property to BFS  $\tilde{A}$  and

$$-1 \leq \mu_{\tilde{A}}^-(x) + \delta_{\tilde{A}}^-(x) \leq 1, \quad (3)$$

for all  $x \in \tilde{X}$ .

**Definition 3** (see [35]). A TFS  $\tilde{A}$  of a non-empty set  $\tilde{X}$  is an object

$$\tilde{A} = \langle \mu_{\tilde{A}}^-, \lambda_{\tilde{A}}^-, \delta_{\tilde{A}}^- \rangle = \left\{ \langle x, \mu_{\tilde{A}}^-(x), \lambda_{\tilde{A}}^-(x), \delta_{\tilde{A}}^-(x) \rangle \mid x \in \tilde{X} \right\}, \quad (4)$$

where  $\mu_{\tilde{A}}^-: \tilde{X} \rightarrow [0, 1]$ ,  $\lambda_{\tilde{A}}^-: \tilde{X} \rightarrow [0, 1]$  and  $\delta_{\tilde{A}}^-: \tilde{X} \rightarrow [-1, 0]$  such that  $0 \leq \mu_{\tilde{A}}^-(x) + \lambda_{\tilde{A}}^-(x) \leq 1$ . The membership degree  $\mu_{\tilde{A}}^-(x)$  is the amount to which the element  $x$  satisfies the condition to the TFS  $\tilde{A}$ ,  $\lambda_{\tilde{A}}^-(x)$  characterizes the extent that the element  $x$  satisfies to the irrelevant property corresponding to tripolar FS  $\tilde{A}$ , and  $\delta_{\tilde{A}}^-(x)$  characterizes the extent that  $x$  satisfies to the implicit counter property of TF set  $\tilde{A}$ . For simplicity, we denote by  $-t = \langle \mu, \lambda, \delta \rangle$  a TFS and call it a TFN.

**Remark 1.** In Definition 3, for a TFS  $-t = \langle \mu, \lambda, \delta \rangle$ , if  $\delta = 0$ , then, it reduces to an IFS  $-t = \langle \mu, \lambda \rangle$ , and if  $\lambda = 0$ , it reduces to a BFS,  $-t = \langle \mu, \delta \rangle$ .

In order to rank TFNs, we need to define the score and accuracy functions.

**Definition 4.** The score function  $\odot(-t)$  of a TFN,  $-t = \langle \mu, \lambda, \delta \rangle$ , is defined as follows:

$$\odot(-t) = \frac{1 + \mu + \lambda + \delta}{3}, \odot(-t) \in [0, 1]. \quad (5)$$

**Definition 5.** The accuracy function  $ac(-t)$  of a TFN  $-t = \langle \mu, \lambda, \delta \rangle$  is evaluated as follows:

$$ac(-t) = \frac{\mu - \delta}{2}, \quad ac(-t) \in [0, 1]. \quad (6)$$

Note that  $ac(-t)$  measures the degree of accuracy of  $-t = \langle \mu, \lambda, \delta \rangle$ . Largest value of  $ac(-t)$  shows that the TFN,  $-t = \langle \mu, \lambda, \delta \rangle$ , is more accurate. In the following, we create an

ordered relationship between two TFNs,  $-t_1 = (\mu_1, \lambda_1, \delta_1)$  and  $-t_2 = (\mu_2, \lambda_2, \delta_2)$ , using  $sc(-t)$  and  $ac(-t)$ .

**Definition 6.** If  $\odot(-t_1) < \odot(-t_2)$  or  $\odot(-t_1) = \odot(-t_2)$  but  $ac(-t_1) < ac(-t_2)$ , then  $-t_1$  is less than  $-t_2$  referred to as  $-t_1 < -t_2$ ; if  $\odot(-t_1) = \odot(-t_2)$  and  $ac(-t_1) < ac(-t_2)$ , then  $-t_1 = -t_2$ .

The following are some basic TFN operations.

**Definition 7.** Let  $-t_r = (\mu_r, \lambda_r, \delta_r)$ , ( $r = 1, 2$ ), and  $-t = (\mu, \lambda, \delta)$  be any three TFNs and  $\kappa > 0$ ; then, we have

- (i)  $-t_1 \oplus -t_2 = (\mu_1 + \mu_2 - \mu_1 \mu_2, \lambda_1 \lambda_2, -|\delta_1| |\delta_2|)$ .
- (ii)  $-t_1 \otimes -t_2 = (\mu_1 \mu_2, \lambda_1 + \lambda_2 - \lambda_1 \lambda_2, \delta_1 + \delta_2 - \delta_1 \delta_2)$ .
- (iii)  $\kappa(-t) = (1 - (1 - \mu)^\kappa, \lambda^\kappa, -|\delta|^\kappa)$ .
- (iv)  $(-t)^\kappa = (\mu^\kappa, 1 - (1 - \lambda)^\kappa, -1 + |1 + \delta|^\kappa)$ .
- (v)  $(-t)^c = (\lambda, \mu, |\delta| - 1)$ .
- (vi)  $-t_1 \subseteq -t_2$  if and only if  $\mu_1 \leq \mu_2$ ,  $\lambda_1 \leq \lambda_2$  and  $\delta_1 \geq \delta_2$ .
- (vii)  $-t_1 \cup -t_2 = (\max\{\mu_1, \mu_2\}, \min\{\lambda_1, \lambda_2\}, \min\{\delta_1, \delta_2\})$ .
- (viii)  $-t_1 \cap -t_2 = (\min\{\mu_1, \mu_2\}, \max\{\lambda_1, \lambda_2\}, \max\{\delta_1, \delta_2\})$ .

We may simply obtain the following operations from Definition 7.

**Theorem 1.** Let  $-t_j = (\mu_r, \lambda_r, \delta_r)$ , ( $r = 1, 2$ ), and  $-t = (\mu, \lambda, \delta)$  be any three TFNs and  $\kappa, \kappa_1, \kappa_2 > 0$ ; then,

- (1)  $-t_1 \oplus -t_2 = -t_2 \oplus -t_1$ .
- (2)  $-t_1 \otimes -t_2 = -t_2 \otimes -t_1$ .
- (3)  $\kappa(-t_1 \oplus -t_2) = \kappa(-t_1) \oplus \kappa(-t_2)$ .
- (4)  $(-t_1 \otimes -t_2)^\kappa = (-t_1)^\kappa \otimes (-t_2)^\kappa$ .
- (5)  $\kappa_1 - t \oplus \kappa_2 - t = (\kappa_1 + \kappa_2) - t$ .
- (6)  $(-t)^{\kappa_1} \otimes (-t)^{\kappa_2} = (-t)^{\kappa_1 + \kappa_2}$ .
- (7)  $((-t)^{\kappa_1})^{\kappa_2} = (-t)^{\kappa_1 \kappa_2}$ .

*Dombi product and Dombi sum are special cases of triangle norms and conorms, which are defined further below.*

**Definition 8** (Dombi [38]). Assume that  $\alpha$  and  $\beta$  can be any two real numbers. The following formulas define Dombi T-norms and T-conorms.

$$\text{Dombi}(\alpha, \beta) = \frac{1}{1 + \left\{ (1 - \alpha/\alpha)^\Re + (1 - \beta/\beta)^\Re \right\}^{(1/\Re)}}, \quad (7)$$

$$\text{Dombi}^c(\alpha, \beta) = 1 - \frac{1}{1 + \left\{ (\alpha/1 - \alpha)^\Re + (\beta/1 - \beta)^\Re \right\}^{(1/\Re)}}, \quad (8)$$

where  $\Re \geq 1$  and  $(\alpha, \beta) \in [0, 1] \times [0, 1]$ .

**2.1. TF-Dombi Operations.** In light of DTN and DTCN, we use TFNs to clarify Dombi operations in this section. On the basis of Dombi operations, we will suggest some TFN operating laws.

**Definition 9.** Let  $-t_r = (\mu_r, \lambda_r, \delta_r)$ , ( $r = 1, 2$ ), be TFNs and  $\mathfrak{R} \geq 1$  and  $\kappa > 0$ ; then, the D T-norm and T-conorm operations of TFNs are introduced as follows:

- (i)  $-t_1 \oplus -t_2 = (1 - \frac{(\mu_1/1 - \mu_1)^{\mathfrak{R}} + (\mu_2/1 - \mu_2)^{\mathfrak{R}}}{(1/1 + \{(\mu_1/1 - \mu_1)^{\mathfrak{R}} + (\mu_2/1 - \mu_2)^{\mathfrak{R}}\}^{(1/\mathfrak{R})})}, \frac{(1/1 + \{(1 - \lambda_1/\lambda_1)^{\mathfrak{R}} + (1 - \lambda_2/\lambda_2)^{\mathfrak{R}}\}^{(1/\mathfrak{R})})}{(-1/1 + \{(\delta_1 + 1/|\delta_1|)^{\mathfrak{R}} + (\delta_2 + 1/|\delta_2|)^{\mathfrak{R}}\}^{(1/\mathfrak{R})})})$ .
- (ii)  $-t_1 \otimes -t_2 = (\frac{(1/1 + \{(1 - \mu_1/\mu_1)^{\mathfrak{R}} + (1 - \mu_2/\mu_2)^{\mathfrak{R}}\}^{(1/\mathfrak{R})})}{\mathfrak{R}\}^{(1/\mathfrak{R})}), 1 - \frac{(1/1 + \{(\lambda_1/1 - \lambda_1)^{\mathfrak{R}} + (\lambda_2/1 - \lambda_2)^{\mathfrak{R}}\}^{(1/\mathfrak{R})})}{-1 + (1/1 + \{(|\delta_1|/\delta_1 + 1)^{\mathfrak{R}} + (|\delta_2|/\delta_2 + 1)^{\mathfrak{R}}\}^{(1/\mathfrak{R})})})$ .
- (iii)  $\kappa(-t_1) = (1 - \frac{(1/1 + \{\kappa(\mu_1/1 - \mu_1)^{\mathfrak{R}}\}^{(1/\mathfrak{R})})}{+ \{\kappa(1 - \lambda_1/\lambda_1)^{\mathfrak{R}}\}^{(1/\mathfrak{R})}}, (-1/1 + \frac{\kappa(\delta_1 + 1/|\delta_1|)^{\mathfrak{R}}}{\mathfrak{R}\}^{(1/\mathfrak{R})})$ .
- (iv)  $(-t_1)^\kappa = (\frac{(1/1 + \{\kappa(1 - \mu_1/\mu_1)^{\mathfrak{R}}\}^{(1/\mathfrak{R})})}{\mathfrak{R}\}^{(1/\mathfrak{R})}), 1 - \frac{(1/1 + \{\kappa(\lambda_1/1 - \lambda_1)^{\mathfrak{R}}\}^{(1/\mathfrak{R})})}{-1 + (1/1 + \{\kappa(|\delta_1|/\delta_1 + 1)^{\mathfrak{R}}\}^{(1/\mathfrak{R})})})$ .

**2.2. TF Dombi Averaging AOs.** In this section, we create and examine the basic features of the TF-Dombi weighted averaging (TFDWA) operator, TF-Dombi ordered weighted averaging (TFDOWA) operator, and TF-Dombi hybrid weight averaging (TFDHWa) operator, which are all arithmetic aggregation operators using TFNs.

**Definition 10.** Let  $-t_r = (\mu_r, \lambda_r, \delta_r)$ , ( $r = 1, \dots, p$ ), be a collection of TFNs. Then, the TF-Dombi TFDWA operator is a mapping TFDWA:  $-t^n \rightarrow -t$  such that

$$\text{TFDWA}_\Phi(-t_1, -t_2, \dots, -t_p) = \bigoplus_{r=1}^p (\Phi_r - t_r), \quad (9)$$

where  $\Phi = (\Phi_1, \Phi_2, \dots, \Phi_p)^T$  is the weight vector of  $-t_r$ , ( $r = 1, \dots, p$ ) with the conditions  $\Phi_r > 0$  and  $\bigoplus_{r=1}^p \Phi_r = 1$ .

**Remark 2.** In Definition 10, the collection of TFNs  $-t_r = (\mu_r, \lambda_r, \delta_r)$ , ( $r = 1, \dots, p$ ), we have

- (i) If  $\delta_r = 0$ , for all  $r$ ,  $1 \leq r \leq p$ , then the family of TFNs,  $-t_r$ , ( $r = 1, \dots, p$ ), reduces to a family of IFNs,  $-t_r = (\mu_r, \lambda_r)$ , ( $r = 1, \dots, p$ ).
- (ii) If  $\lambda_r = 0$ , for all  $r$ ,  $1 \leq r \leq p$ , then the family of TFNs,  $-t_r$ , ( $r = 1, \dots, p$ ), reduces to a family of BFNs,  $-t_r = (\mu_r, \delta_r)$ , ( $r = 1, \dots, p$ ).

In the theorem mentioned below, we use the Dombi operation on TFNs and develop TFDWA operator for the aggregation of TFNs.

**Theorem 2.** Let  $-t_r = (\mu_r, \lambda_r, \delta_r)$ , ( $r = 1, \dots, p$ ), be a family of TFNs; then, the aggregated value of this family by using the TFDWA operator is also a TFN, and

$$\begin{aligned} & \text{TFDWA}_\Phi(-t_1, -t_2, \dots, -t_p) \\ &= \bigoplus_{r=1}^p (\Phi_r - t_r) = \left( \frac{1 - \frac{1}{1 + \{\bigoplus_{r=1}^p \Phi_r (\mu_r/1 - \mu_r)^{\mathfrak{R}}\}^{(1/\mathfrak{R})}}, \frac{1}{1 + \{\bigoplus_{r=1}^p \Phi_r (1 - \lambda_r/\lambda_r)^{\mathfrak{R}}\}^{(1/\mathfrak{R})}}}{-1} \right. \\ & \left. \frac{-1}{1 + \{\bigoplus_{r=1}^p \Phi_r (1 + \delta_r/|\delta_r|)^{\mathfrak{R}}\}^{(1/\mathfrak{R})}} \right), \end{aligned} \quad (10)$$

where  $\Phi = (\Phi_1, \Phi_2, \dots, \Phi_p)^T$  is the weight vector of  $-t_r$ , ( $r = 1, \dots, p$ ) with  $\Phi_r > 0$  and  $\bigoplus_{r=1}^p \Phi_r = 1$ .

*Proof.* The mathematical induction approach can be used to prove this theorem. Thus,

- (i) When  $p = 2$ , based on TFNs' Dombi operation, as defined in Definition 9, we obtain

$$\text{TFDWA}_\Phi(-t_1, -t_2) = \Phi_1 - t_1 \oplus \Phi_2 - t_2 = \Phi_1 (\mu_1, \lambda_1, \delta_1) \oplus \Phi_2 (\mu_2, \lambda_2, \delta_2). \quad (11)$$

For the RHS of equation (10), we have

$$\begin{aligned}
 & \left( \frac{1 - \frac{1}{1 + \{\Phi_1(\mu_1/1 - \mu_1)^{\Re} + \Phi_2(\mu_2/1 - \mu_2)^{\Re}\}^{(1/\Re)}}}}{1 + \{\Phi_1(1 - \lambda_1/\lambda_1)^{\Re} + \Phi_2(1 - \lambda_2/\lambda_2)^{\Re}\}^{(1/\Re)}}}, \frac{1}{1 + \{\Phi_1(1 - \lambda_1/\lambda_1)^{\Re} + \Phi_2(1 - \lambda_2/\lambda_2)^{\Re}\}^{(1/\Re)}}} \right) \\
 & \left( \frac{-1}{1 + \{\Phi_1(1 + \delta_1/|\delta_1|)^{\Re} + \Phi_2(1 + \delta_2/|\delta_2|)^{\Re}\}^{(1/\Re)}}} \right) \\
 & = \left( \frac{1 - \frac{1}{1 + \{\sum_{r=1}^2 \Phi_r(\mu_r/1 - \mu_r)^{\Re}\}^{(1/\Re)}}}}{1 + \{\sum_{r=1}^2 \Phi_r(1 - \lambda_r/\lambda_r)^{\Re}\}^{(1/\Re)}}}, \frac{1}{1 + \{\sum_{r=1}^2 \Phi_r(1 - \lambda_r/\lambda_r)^{\Re}\}^{(1/\Re)}}} \right) \\
 & \left( \frac{-1}{1 + \{\sum_{r=1}^2 \Phi_r(1 + \delta_r/|\delta_r|)^{\Re}\}^{(1/\Re)}}} \right).
 \end{aligned} \tag{12}$$

Thus, equation (9) holds for  $p = 2$ .

(ii) Assume that equation (9) holds for  $p = k$ , where  $k \in \mathbb{N}$  (set of natural numbers); then, equation (9) becomes

$$\begin{aligned}
 & \text{TFDWA}_{\Phi}(\xi_1, \xi_2, \dots, \xi_k) \\
 & = \bigoplus_{r=1}^k (\Phi_r - t_r) = \left( \frac{1 - \frac{1}{1 + \{\sum_{r=1}^k \Phi_r(\mu_r/1 - \mu_r)^{\Re}\}^{(1/\Re)}}}}{1 + \{\sum_{r=1}^k \Phi_r(1 - \lambda_r/\lambda_r)^{\Re}\}^{(1/\Re)}}}, \frac{1}{1 + \{\sum_{r=1}^k \Phi_r(1 - \lambda_r/\lambda_r)^{\Re}\}^{(1/\Re)}}} \right) \\
 & \left( \frac{-1}{1 + \{\sum_{r=1}^k \Phi_r(1 + \delta_r/|\delta_r|)^{\Re}\}^{(1/\Re)}}} \right).
 \end{aligned} \tag{13}$$

Now for  $p = k + 1$ , we have the following equation.

$$\begin{aligned}
 & \text{TFDWA}_{\Phi}(-t_1, -t_2, \dots, -t_k, -t_{k+1}) \\
 & = \bigoplus_{r=1}^k (\Phi_r - t_r) \oplus (\Phi_{k+1} - t_{k+1}) \\
 & = \left( \frac{1 - \frac{1}{1 + \{\sum_{r=1}^k \Phi_r(\mu_r/1 - \mu_r)^{\Re}\}^{(1/\Re)}}}}{1 + \{\sum_{r=1}^k \Phi_r(1 - \lambda_r/\lambda_r)^{\Re}\}^{(1/\Re)}}}, \frac{1}{1 + \{\sum_{r=1}^k \Phi_r(1 - \lambda_r/\lambda_r)^{\Re}\}^{(1/\Re)}}} \right) \\
 & \left( \frac{-1}{1 + \{\sum_{r=1}^k \Phi_r(1 + \delta_r/|\delta_r|)^{\Re}\}^{(1/\Re)}}} \right) \\
 & \oplus \left( \frac{1 - \frac{1}{1 + \{\Phi_{k+1}(\mu_{k+1}/1 - \mu_{k+1})^{\Re}\}^{(1/\Re)}}}}{1 + \{\Phi_{k+1}(1 - \lambda_{k+1}/\lambda_{k+1})^{\Re}\}^{(1/\Re)}}}, \frac{1}{1 + \{\Phi_{k+1}(1 - \lambda_{k+1}/\lambda_{k+1})^{\Re}\}^{(1/\Re)}}} \right) \\
 & \left( \frac{-1}{1 + \{\Phi_{k+1}(1 + \delta_{k+1}/|\delta_{k+1}|)^{\Re}\}^{(1/\Re)}}} \right) \\
 & = \left( \frac{1 - \frac{1}{1 + \{\sum_{r=1}^{k+1} \Phi_r(\mu_r/1 - \mu_r)^{\Re}\}^{(1/\Re)}}}}{1 + \{\sum_{r=1}^{k+1} \Phi_r(1 - \lambda_r/\lambda_r)^{\Re}\}^{(1/\Re)}}}, \frac{1}{1 + \{\sum_{r=1}^{k+1} \Phi_r(1 - \lambda_r/\lambda_r)^{\Re}\}^{(1/\Re)}}} \right) \\
 & \left( \frac{-1}{1 + \{\sum_{r=1}^{k+1} \Phi_r(1 + \delta_r/|\delta_r|)^{\Re}\}^{(1/\Re)}}} \right).
 \end{aligned} \tag{14}$$

Thus, equation (9) is true for  $p = k + 1$ . Therefore, equation (9) is true for any  $p \in \mathbb{N}$ .  $\square$

*Remark 3.* In Theorem 2, as part of a collection of TFNs,  $-t_r = (\mu_r, \lambda_r, \delta_r)$ , ( $r = 1, \dots, p$ ), if

$$\begin{aligned} & \text{IFDWA}_\Phi(-t_1, -t_2, \dots, -t_p) \\ &= \bigoplus_{r=1}^p (\Phi_r - t_r) = \left( 1 - \frac{1}{1 + \{\bigoplus_{r=1}^p \Phi_r (\mu_r/1 - \mu_r)\}^{\mathfrak{R}} (1/\mathfrak{R})}, \frac{1}{1 + \{\bigoplus_{r=1}^p \Phi_r (1 - \lambda_r/\lambda_r)\}^{\mathfrak{R}} (1/\mathfrak{R})} \right). \end{aligned} \quad (15)$$

(ii) If  $\lambda_r = 0$ , for all  $r$ , then the collection of TFNs,  $-t_r = (\mu_r, \lambda_r, \delta_r)$ , becomes a collection of BFNs,  $-t_r = (\mu_r, \delta_r)$ , ( $r = 1, \dots, p$ ), and the TFDWA

(i)  $\delta_r = 0$ , for all  $r$ , then the collection of TFNs,  $-t_r = (\mu_r, \lambda_r, \delta_r)$ , becomes a collection of IFNs,  $-t_r = (\mu_r, \lambda_r)$ , ( $r = 1, \dots, p$ ), and the TFDWA operator reduces to the IFDWA operator, as shown below.

operator reduces to the BFDWA operator, as given below.

$$\begin{aligned} & \text{BFDWA}_\Phi(-t_1, -t_2, \dots, -t_p) \\ &= \bigoplus_{r=1}^p (\Phi_r - t_r) = \left( 1 - \frac{1}{1 + \{\bigoplus_{r=1}^p \Phi_r (\mu_r/1 - \mu_r)\}^{\mathfrak{R}} (1/\mathfrak{R})}, \frac{-1}{1 + \{\bigoplus_{r=1}^p \Phi_r (1 + \delta_r/|\delta_r|)\}^{\mathfrak{R}} (1/\mathfrak{R})} \right). \end{aligned} \quad (16)$$

The TFDWA operator has a number of essential qualities, which are given below.

**Theorem 3 (idempotency property).** If  $-t_r = (\mu_r, \lambda_r, \delta_r)$ , ( $r = 1, \dots, p$ ), is a set of TFNs that all have the same value, i.e., if  $-t_r = (\mu_r, \lambda_r, \delta_r) = -t = (\mu, \lambda, \delta)$  for all  $r$ , then

$$\text{TFDWA}_\Phi(-t_1, -t_2, \dots, -t_r) = -t. \quad (17)$$

*Proof.* Since  $-t_r = (\mu_r, \lambda_r, \delta_r) = -t$  ( $r = 1, \dots, p$ ), then by using equation (9), we have

$$\begin{aligned} & \text{TFDWA}_\Phi(-t_1, -t_2, \dots, -t_r) \\ &= \bigoplus_{r=1}^p (\Phi_r - t_r) = \left( \frac{1 - \frac{1}{1 + \{\bigoplus_{r=1}^p \Phi_r (\mu_r/1 - \mu_r)\}^{\mathfrak{R}} (1/\mathfrak{R})}}{1 + \{\bigoplus_{r=1}^p \Phi_r (1 - \lambda_r/\lambda_r)\}^{\mathfrak{R}} (1/\mathfrak{R})}, \frac{-1}{1 + \{\bigoplus_{r=1}^p \Phi_r (1 + \delta_r/|\delta_r|)\}^{\mathfrak{R}} (1/\mathfrak{R})} \right) \\ &= \left( 1 - \frac{1}{1 + \{(\mu/1 - \mu)\}^{\mathfrak{R}} (1/\mathfrak{R})}, \frac{1}{1 + \{(1 - \lambda/\lambda)\}^{\mathfrak{R}} (1/\mathfrak{R})}, \frac{-1}{1 + \{(1 + \delta/|\delta|)\}^{\mathfrak{R}} (1/\mathfrak{R})} \right) \\ &= \left( 1 - \frac{1}{1 + (\mu/1 - \mu)}, \frac{1}{1 + (1 - \lambda/\lambda)}, \frac{-1}{1 + (1 + \delta/|\delta|)} \right) = (\mu, \lambda, \delta). \end{aligned} \quad (18)$$

Thus,  $\text{TFDWA}_\Phi(-t_1, -t_2, \dots, -t_r) = t$ .

In a similar manner, we can verify the following properties of the TFDWA operator.  $\square$

**Theorem 4 (boundedness property).** Let  $-t_r = (\mu_r, \lambda_r, \delta_r)$ , ( $r = 1, 2, \dots, p$ ), be a family of TFNs, and  $-t^- = \min\{-t_r\}$  and  $-t^+ = \max\{-t_r\}$ . Then,

$$-t^- \leq \text{TFDWA}_\Phi(-t_1, -t_2, \dots, -t_p) \leq -t^+. \quad (19)$$

**Theorem 5 (monotonicity property).** If  $-t_r = (\mu_r, \lambda_r, \delta_r)$ ,  $(r = 1, \dots, p)$ , and  $-t'_r = (\mu'_r, \lambda'_r, \delta'_r)$ ,  $(r = 1, 2, \dots, p)$ , are the families of TFNs, such that  $-t_r \leq -t'_r$  for all  $r$ , then

$$\text{TFDWA}_\Phi(-t_1, -t_2, \dots, -t_p) \leq \text{TFDWA}_\Phi(-t'_1, -t'_2, \dots, -t'_p). \quad (20)$$

Now we will go through the fundamental features of the TFDWA operator.

**Definition 11.** Let  $-t_r = (\mu_r, \lambda_r, \delta_r)$ ,  $(r = 1, \dots, p)$ , be a collection of TFNs. A TFDWA operator of dimension  $p$  is a mapping  $\text{TFDWA}_\omega: -t^p \rightarrow t$  with an associated weight

vector  $\omega = (\omega_1, \omega_2, \dots, \omega_p)^T$  such that  $\omega_r > 0$ , and  $\oplus_{r=1}^p \omega_r = 1$ . Therefore,

$$\text{TFDWA}_\Phi(-t_1, -t_2, \dots, -t_p) = \oplus_{r=1}^p (\Phi_r - t_{r(\sigma)}), \quad (21)$$

where  $(\sigma(1), \sigma(2), \dots, \sigma(p))$  are the permutations of  $\sigma(r)$   $(r = 1, 2, \dots, p)$  for which  $-t_{\sigma(r-1)} \geq -t_{\sigma(r)}$  for all  $r = 1, 2, \dots, p$ .

The following theorem is a consequence of Definition 10 and Theorem 2.

**Theorem 6.** Let  $-t_r = (\mu_r, \lambda_r, \delta_r)$ ,  $(r = 1, \dots, p)$ , be a family of TFNs. A TFDWA operator of dimension  $p$  is a mapping from  $-t^p$  to  $t$  with the associated weight vector  $\omega = (\omega_1, \omega_2, \dots, \omega_p)^T$  such that  $\omega_r > 0$  and  $\oplus_{r=1}^p \omega_r = 1$ . Then,

$$\begin{aligned} & \text{TFDWA}_\omega(-t_1, -t_2, \dots, -t_p) \\ &= \oplus_{r=1}^p (\Phi_r - t_{r(\sigma)}) = \left( \frac{1 - \frac{1}{1 + \left\{ \oplus_{r=1}^p \omega_r (\mu_{\sigma(r)} / 1 - \mu_{\sigma(r)})^{\Re} \right\}^{(1/\Re)}}}}{1 + \left\{ \oplus_{r=1}^p \omega_r (1 - \lambda_{\sigma(r)} / \lambda_{\sigma(r)})^{\Re} \right\}^{(1/\Re)}}, \frac{1}{1 + \left\{ \oplus_{r=1}^p \omega_r (1 + \delta_{\sigma(r)} / |\delta_{\sigma(r)}|)^{\Re} \right\}^{(1/\Re)}}} \right), \end{aligned} \quad (22)$$

where  $(\sigma(1), \sigma(2), \dots, \sigma(p))$  are the permutations of  $\sigma(r)$   $(r = 1, 2, \dots, p)$  for which  $-t_{\sigma(r-1)} \geq -t_{\sigma(r)}$  for all  $r = 1, 2, \dots, p$ .

**Remark 4.** In Definition 10, the collection of TFNs  $-t_r = (\mu_r, \lambda_r, \delta_r)$ ,  $(r = 1, \dots, p)$ , we have

- (i) If  $\delta_r = 0$ , for all  $r$ , then the TFDWA operator decreases to IFDWA operator and

$$\begin{aligned} & \text{IFDWA}_\omega(-t_1, -t_2, \dots, -t_p) \\ &= \oplus_{r=1}^p (\Phi_r - t_{r(\sigma)}) = \left( 1 - \frac{1}{1 + \left\{ \oplus_{r=1}^p \omega_r (\mu_{\sigma(r)} / 1 - \mu_{\sigma(r)})^{\Re} \right\}^{(1/\Re)}}, \frac{1}{1 + \left\{ \oplus_{r=1}^p \omega_r (1 - \lambda_{\sigma(r)} / \lambda_{\sigma(r)})^{\Re} \right\}^{(1/\Re)}}} \right). \end{aligned} \quad (23)$$

- (ii) If  $\lambda_r = 0$ , for all  $r$ , then the TFDWA operator decreases to BFDWA operator and

$$\begin{aligned} & \text{BFDWA}_\omega(-t_1, -t_2, \dots, -t_p) \\ &= \oplus_{r=1}^p (\Phi_r - t_{r(\sigma)}) = \left( 1 - \frac{1}{1 + \left\{ \oplus_{r=1}^p \omega_r (\mu_{\sigma(r)} / 1 - \mu_{\sigma(r)})^{\Re} \right\}^{(1/\Re)}}, \frac{1}{1 + \left\{ \oplus_{r=1}^p \omega_r (1 + \delta_{\sigma(r)} / |\delta_{\sigma(r)}|)^{\Re} \right\}^{(1/\Re)}}} \right). \end{aligned} \quad (24)$$



The following properties of a TFDOWA operator can easily be proved.

(P1) (idempotency property). If  $-t_r = (\mu_r, \lambda_r, \delta_r)$  ( $r = 1, 2, \dots, p$ ) are all equal, i.e.,  $-t_r = -t$  for all  $r$ , then  $\text{TFDOWA}_\omega(-t_1, -t_2, \dots, -t_p) = t$ .

(P2) (boundedness property). Let  $-t_r = (\mu_r, \lambda_r, \delta_r)$ , ( $r = 1, 2, \dots, p$ ), be a family of TFNs and  $-t^- = \min -t_r$  and  $-t^+ = \max -t_r$ . Then,

$$-t^- \leq \text{TFDOWA}_\omega(-t_1, -t_2, \dots, -t_p) \leq -t^+. \quad (25)$$

(P3) (monotonicity property). If  $-t_r = (\mu_r, \lambda_r, \delta_r)$ , ( $r = 1, 2, \dots, p$ ), and  $-t'_r = (\mu'_r, \lambda'_r, \delta'_r)$ , ( $r = 1, 2, \dots, p$ ), are two TFNs such that  $-t_r \leq -t'_r$  for all  $r$ , then

$$\text{TFDOWA}_\omega(-t_1, -t_2, \dots, -t_p) \leq \text{TFDOWA}_\omega(-t'_1, -t'_2, \dots, -t'_p). \quad (26)$$

*Definition 12.* A TFDHWA operator of dimension  $p$  is a mapping  $\text{TFDHA}: -t^p \rightarrow t$  with the associated weight vector  $w = (w_1, w_2, \dots, w_p)$  such that  $w_r > 0$  and  $\oplus_{r=1}^p w_r = 1$

and the TFDHWA operator is provided by the following equation:

$$\begin{aligned} & \text{TFDHWA}_{(w, \Phi)}(-t_1, -t_2, \dots, -t_p) \\ &= ((w_1 - t_{\sigma(1)}), (w_2 - t_{\sigma(2)}), \dots, (w_p - t_{\sigma(p)})) \\ &= \left( \frac{1 - \frac{1}{1 + \left\{ \oplus_{r=1}^p w_r (\mu_{\sigma(r)} / 1 - \mu_{\sigma(r)})^{\mathfrak{R}} \right\}^{(1/\mathfrak{R})}}}}{1 + \left\{ \oplus_{r=1}^p w_r (1 - \lambda_{\sigma(r)} / \lambda_{\sigma(r)})^{\mathfrak{R}} \right\}^{(1/\mathfrak{R})}}, \frac{1}{1 + \left\{ \oplus_{r=1}^p w_r (1 + \delta_{\sigma(r)} / |\delta_{\sigma(r)}|)^{\mathfrak{R}} \right\}^{(1/\mathfrak{R})}} \right), \end{aligned} \quad (27)$$

where  $-t_{\sigma(r)}$  is the  $r^{\text{th}}$  largest WTF number  $-t_r$ , ( $-t_r = p\Phi_r - t_r$ ,  $r = 1, 2, \dots, p$ ) and  $\Phi = (\Phi_1, \Phi_2, \dots, \Phi_p)^T$  is the weight vector of  $-t_r$  with the condition  $\Phi_r > 0$  and  $\oplus_{r=1}^p \Phi_r = 1$ , where  $p$  is the balancing coefficient. Some important properties of (TFDHWA) operator are given below:

(i) When  $w = ((1/p), (1/p), \dots, (1/p))$ , then TFDWA operator is a special case of TFDHA operator.

(ii) When,  $\Phi = ((1/p), (1/p), \dots, (1/p))$ , then TFDOWA operator is a special case of the TFDHWA operator. As a result, we conclude that the TFDHWA operator is a generalization of both the TFDWA and TFDOWA operators.

*Remark 5.* In Definition 12, the collection of TFNs  $-t_r = (\mu_r, \lambda_r, \delta_r)$ , ( $r = 1, \dots, p$ ), we have

- (i) If  $\delta_r = 0$ , for all  $r$ , then the TFDHWA operator reduces to IFDHWA operator and

$$\begin{aligned}
 & \text{IFDHWA}_{(w,\Phi)}(-t_1, -t_2, \dots, -t_p) \\
 &= ((w_1 - t_{\sigma(1)}), (w_2 - t_{\sigma(2)}), \dots, (w_p - t_{\sigma(p)})) \\
 &= \left( 1 - \frac{1}{1 + \left\{ \oplus_{r=1}^p w_r (\mu_{\sigma(r)}/1 - \mu_{\sigma(r)})^{\mathfrak{R}} \right\}^{(1/\mathfrak{R})}}, \frac{1}{1 + \left\{ \oplus_{r=1}^p w_r (1 - \lambda_{\sigma(r)}/\lambda_{\sigma(r)})^{\mathfrak{R}} \right\}^{(1/\mathfrak{R})}} \right).
 \end{aligned} \tag{28}$$

- (ii) If  $\lambda_r = 0$ , for all  $r$ , then the TFDHWA operator reduces to BFDHWA operator and

$$\begin{aligned}
 & \text{BFDHWA}_{(w,\Phi)}(-t_1, -t_2, \dots, -t_p) \\
 &= ((w_1 - t_{\sigma(1)}), (w_2 - t_{\sigma(2)}), \dots, (w_p - t_{\sigma(p)})) \\
 &= \left( 1 - \frac{1}{1 + \left\{ \oplus_{r=1}^p w_r (\mu_{\sigma(r)}/1 - \mu_{\sigma(r)})^{\mathfrak{R}} \right\}^{(1/\mathfrak{R})}}, \frac{1}{1 + \left\{ \oplus_{r=1}^p w_r (1 + \delta_{\sigma(r)}/|\delta_{\sigma(r)}|)^{\mathfrak{R}} \right\}^{(1/\mathfrak{R})}} \right).
 \end{aligned} \tag{29}$$

We now, give an illustrative example for the verification of Definition 13.

*Example 1.* Consider four TFNs,  $-t_1 = (0.5, 0.4, -0.3)$ ,  $-t_2 = (0.6, 0.4, -0.3)$ ,  $-t_3 = (0.7, 0.2, -0.3)$ , and

$-t_4 = (0.2, 0.4, -0.4)$ , and let  $\Phi = (0.20, 0.30, 0.30, 0.20)^T$  be the associated weight vector. Then, by Definition 13, the aggregated value of TFNs for  $\mathfrak{R} = 3$ , and by using the TFDHWA operator, we get

$$\begin{aligned}
 -t_1 &= \left( \frac{1 - \frac{1}{1 + \left\{ 4 \times 0.20 (0.5/1 - 0.5)^3 \right\}^{(1/3)}}, \frac{1}{1 + \left\{ 4 \times 0.20 (1 - 0.4/0.4)^3 \right\}^{(1/3)}}}{\frac{-1}{1 + \left\{ 4 \times 0.20 (1 - 0.3/0.3)^3 \right\}^{(1/3)}}} \right) \\
 &= (0.4814, 0.4179, -0.3158) \\
 -t_2 &= \left( \frac{1 - \frac{1}{1 + \left\{ 4 \times 0.30 (0.6/1 - 0.6)^3 \right\}^{(1/3)}}, \frac{1}{1 + \left\{ 4 \times 0.30 (1 - 0.4/0.4)^3 \right\}^{(1/3)}}}{\frac{-1}{1 + \left\{ 4 \times 0.30 (1 - 0.3/0.3)^3 \right\}^{(1/3)}}} \right) \\
 &= (0.6144, 0.3855, -0.2873)
 \end{aligned}$$

$$\begin{aligned}
-t_3 &= \left( \begin{array}{c} 1 - \frac{1}{1 + \{4 \times 0.30 (0.7/1 - 0.7)^3\}^{(1/3)}} \frac{1}{1 + \{4 \times 0.30 (1 - 0.2/0.2)^3\}^{(1/3)}} \\ \frac{-1}{1 + \{4 \times 0.30 (1 - 0.3/0.3)^3\}^{(1/3)}} \end{array} \right) = (0.7126, 0.1904, -0.2874) - t_4 \\
&= \left( \begin{array}{c} 1 - \frac{1}{1 + \{4 \times 0.20 (0.2/1 - 0.2)^3\}^{(1/3)}} \frac{1}{1 + \{4 \times 0.20 (1 - 0.4/0.4)^3\}^{(1/3)}} \\ \frac{-1}{1 + \{4 \times 0.20 (1 - 0.4/0.4)^3\}^{(1/3)}} \end{array} \right) = (0.1883, 0.4179, -0.4179). \quad (30)
\end{aligned}$$

Scores of  $-t_r$ , ( $r = 1, 2, 3, 4$ ) are calculated as follows:

$$\begin{aligned}
\textcircled{\text{S}}(-t_1) &= (1 + 0.4814 + 0.4179 - 0.3158/3) = 0.5278, \\
\textcircled{\text{S}}(-t_2) &= (1 + 0.6144 + 0.3855 - 0.2873/3) = 0.5708, \\
\textcircled{\text{S}}(-t_3) &= (1 + 0.7126 + 0.1904 - 0.2874/3) = 0.5385, \\
\textcircled{\text{S}}(-t_4) &= (1 + 0.1883 + 0.3179 - 0.4179/3) = 0.3961.
\end{aligned}$$

Since  $\textcircled{\text{S}}(-t_2) > \textcircled{\text{S}}(-t_3) > \textcircled{\text{S}}(-t_1) > \textcircled{\text{S}}(-t_4)$ , then  $-t_{\sigma(1)} = -t_2 = (0.6145, 0.3855, -0.3158)$ ,  $-t_{\sigma(2)} = -t_3 = (0.7126, 0.1904, -0.2874)$ ,  $-t_{\sigma(3)} = -t_1 = (0.6145, 0.3179, -0.2874)$ , and  $-t_{\sigma(4)} = -t_4 = (0.1884, 0.3419, -0.4180)$ . Therefore, by using TFDHWA operator for  $\mathfrak{R} = 3$ , we get

$$\begin{aligned}
&\text{TFDHWA}(-t_{\sigma(1)}, -t_{\sigma(2)}, \dots, -t_{\sigma(4)}) \\
&= \textcircled{\text{S}}_{r=1}^4(\Phi_r - t_{\sigma(r)}) \\
&= \left( \begin{array}{c} 1 - \frac{1}{1 + \left\{ \sum_{r=1}^4 \Phi_r (\mu_{\sigma(r)}/1 - \mu_{\sigma(r)})^{\mathfrak{R}} \right\}^{(1/\mathfrak{R})}} \frac{1}{1 + \left\{ \sum_{r=1}^4 \Phi_r (1 - \lambda_{\sigma(r)}/\lambda_{\sigma(r)})^{\mathfrak{R}} \right\}^{(1/\mathfrak{R})}} \\ \frac{-1}{1 + \left\{ \sum_{r=1}^4 \Phi_r (1 + \delta_{\sigma(r)}/|\delta_{\sigma(r)}|)^{\mathfrak{R}} \right\}^{(1/\mathfrak{R})}} \end{array} \right) \quad (31) \\
&= \left( \begin{array}{c} 1 - \frac{1}{1 + \{0.2(0.6145/1 - 0.6145)^3 + 0.3(0.7126/1 - 0.7126)^3 + 0.3(0.6145/1 - 0.6145)^3 + 0.2(0.1884/1 - 0.1884)^3\}^{(1/3)}} \\ \frac{1}{1 + \{0.2(1 - 0.3855/0.3855)^3 + 0.3(1 - 0.1904/0.3179)^3 + 0.3(1 - 0.3179/0.3179)^3 + 0.2(1 - 0.3419/0.3419)^3\}^{(1/3)}} \\ \frac{-1}{1 + \{0.2(1 - 0.3158/0.3158)^3 + 0.3(1 - 0.2874/0.2874)^3 + 0.3(1 - 0.2874/0.2874)^3 + 0.2(1 - 0.4180/0.4180)^3\}^{(1/3)}} \end{array} \right) \\
&= (0.1831, 0.1890, -0.2602).
\end{aligned}$$

**2.3. TF Dombi Geometric Aggregation Operators.** In this section, we will propose and examine several Dombi geometric AOs, namely, the TFDWG operator and the TFDOWG operator, utilizing TFNs.

**Definition 13.** Let  $-t_r = (\mu_r, \lambda_r, \delta_r)$ , ( $r = 1, \dots, p$ ), be a family of TFNs. The TFDWG operator is a representation  $-t^p \rightarrow -t$  such that

$$\text{TFDWG}_{\Phi}(-t_1, -t_2, \dots, -t_p) = \textcircled{\text{S}}_{r=1}^p(-t_r)^{\Phi_r}, \quad (32)$$

where  $\Phi = (\Phi_1, \Phi_2, \dots, \Phi_p)^T$  is the weight vector of  $-t_r$ , ( $r = 1, 2, \dots, p$ ) such that  $\Phi_r > 0$  and  $\sum_{r=1}^p \Phi_r = 1$ .

We can easily derive the following theorem from Definition 14.

**Theorem 7.** Let  $-t_r = (\mu_r, \lambda_r, \delta_r)$ , ( $r = 1, 2, \dots, p$ ) be a collection of TFNs. Then, the aggregated value of TFNs is again a TFN, and by using TFDWG operator, we obtain

$$\begin{aligned} & \text{TFDWG}_\Phi(-t_1, -t_2, \dots, -t_p) \\ &= \otimes_{r=1}^p (-t_r)^{\Phi_r} = \left( \frac{1}{1 + \{\oplus_{r=1}^p \Phi_r (1 - \mu_r / \mu_r)\}^{\mathfrak{R}} (1/\mathfrak{R})}, 1 - \frac{1}{1 + \{\oplus_{r=1}^p \Phi_r (\lambda_r / 1 - \lambda_r)\}^{\mathfrak{R}} (1/\mathfrak{R})}, \right. \\ & \quad \left. -1 + \frac{1}{1 + \{\oplus_{r=1}^p \Phi_r (|\delta_r| / 1 + \delta_r)\}^{\mathfrak{R}} (1/\mathfrak{R})} \right), \end{aligned} \quad (33)$$

where  $\Phi = (\Phi_1, \Phi_2, \dots, \Phi_p)^T$  is the weight vector of  $-t_r$ , ( $r = 1, 2, \dots, p$ ) such that  $\Phi_r > 0$  and  $\oplus_{r=1}^p \Phi_r = 1$ .

*Proof.* By using mathematical induction, the proof is straightforward.  $\square$

*Remark 6.* In Theorem 7,

(i) If  $\delta_r = 0$ , for all  $r$ , then the (TFDWG) operator reduces to the IFDWG operator, as given below.

$$\begin{aligned} & \text{IFDWG}_\Phi(-t_1, -t_2, \dots, -t_p) \\ &= \otimes_{r=1}^p (-t_r)^{\Phi_r} = \left( \frac{1}{1 + \{\oplus_{r=1}^p \Phi_r (1 - \mu_r / \mu_r)\}^{\mathfrak{R}} (1/\mathfrak{R})}, 1 - \frac{1}{1 + \{\oplus_{r=1}^p \Phi_r (\lambda_r / 1 - \lambda_r)\}^{\mathfrak{R}} (1/\mathfrak{R})} \right). \end{aligned} \quad (34)$$

(ii) If  $\lambda_r = 0$ , for all  $r$ , then the TFDWG operator reduces to the BFDWG operator, as given below.

$$\begin{aligned} & \text{BFDWG}_\Phi(-t_1, -t_2, \dots, -t_p) \\ &= \otimes_{r=1}^p (-t_r)^{\Phi_r} = \left( \frac{1}{1 + \{\oplus_{r=1}^p \Phi_r (1 - \mu_r / \mu_r)\}^{\mathfrak{R}} (1/\mathfrak{R})}, -1 + \frac{1}{1 + \{\oplus_{r=1}^p \Phi_r (|\delta_r| / 1 + \delta_r)\}^{\mathfrak{R}} (1/\mathfrak{R})} \right). \end{aligned} \quad (35)$$

TFDWG operators have some properties which are listed below:

P1 (idempotency property). If  $-t_r = (\mu_r, \lambda_r, \delta_r)$ , ( $r = 1, \dots, p$ ), is a family of TFNs so that they are all equal, i.e., if  $-t_r = (\mu_r, \lambda_r, \delta_r) = -t = (\mu, \lambda, \delta)$  for all  $r$ , then

$$\text{TFDWG}_\Phi(-t_1, -t_2, \dots, -t_r) = -t. \quad (36)$$

P2 (boundedness property). Let  $-t_r = (\mu_r, \lambda_r, \delta_r)$ , ( $r = 1, 2, 3 \dots, p$ ) be a family of TFNs and  $-t^- = \min -t_r$  and  $-t^+ = \max -t_r$ . Then,

$$-t^- \leq \text{TFDWG}_\Phi(-t_1, -t_2, \dots, -t_p) \leq -t^+. \quad (37)$$

P3 (monotonicity property). If  $-t_r = (\mu_r, \lambda_r, \delta_r)$ , ( $r = 1, 2, \dots, p$ ), and  $-t'_r = (\mu'_r, \lambda'_r, \delta'_r)$ , ( $r = 1, 2, \dots, p$ ), are two TFNs such that  $-t_r \leq -t'_r$  for all  $r$ , then

$$\text{TFDWG}_\Phi(-t_1, -t_2, \dots, -t_p) \leq \text{TFDWG}_\Phi(-t'_1, -t'_2, \dots, -t'_p). \quad (38)$$

The next section introduces the TFDOWG operator.

*Definition 14.* Let  $-t_r = (\mu_r, \lambda_r, \delta_r)$ , ( $r = 1, 2, 3 \dots, p$ ), be a family of TFNs. The TFDOWG operator is a mapping  $-t^p \rightarrow -t$  such that

$$\text{TFDOWG}_w(-t_1, -t_2, \dots, -t_p) = \otimes_{r=1}^p (-t_{\sigma(r)})^{w_r}, \quad (39)$$

where  $w = (w_1, w_2, \dots, w_p)^T$  is an associated weight vector of  $-t_r$ , ( $r = 1, 2, \dots, p$ ) such that  $w_r > 0$  and  $\oplus_{r=1}^p w_r = 1$  and  $(\sigma(1), \sigma(2), \dots, \sigma(p))$  are the permutations of  $-t_r$ , ( $r = 1, 2, \dots, p$ ) for which  $-t_{\sigma(r-1)} \geq -t_{\sigma(r)}$  for all  $r = 1, 2, \dots, p$ .

Using TFDOWG operator, we have the theorem mentioned below.

**Theorem 8.** Let  $-t_r = (\mu_r, \lambda_r, \delta_r)$ ,  $(r = 1, 2, \dots, p)$ , be a family of TFNs. The TFDOWG operator is a mapping  $-t^p \rightarrow -t$  such that

$$\begin{aligned} & \text{TFDOWG}_w(-t_1, -t_2, \dots, -t_p) \\ &= \otimes_{r=1}^p (-t_{\sigma(r)})^{w_r} = \left( \begin{array}{l} \frac{1}{1 + \left\{ \oplus_{r=1}^p \Phi_r(1 - \mu_{\sigma(r)}/\mu_{\sigma(r)})^{\mathfrak{R}} \right\}^{(1/\mathfrak{R})}}, 1 - \frac{1}{1 + \left\{ \oplus_{r=1}^p \Phi_r(\lambda_{\sigma(r)}/1 - \lambda_{\sigma(r)})^{\mathfrak{R}} \right\}^{(1/\mathfrak{R})}}, \\ -1 + \frac{1}{1 + \left\{ \oplus_{r=1}^p \Phi_r(|\delta_{\sigma(r)}|/1 + \delta_{\sigma(r)})^{\mathfrak{R}} \right\}^{(1/\mathfrak{R})}} \end{array} \right), \end{aligned} \quad (40)$$

where  $w = (w_1, w_2, \dots, w_p)^T$  is an associated weight vector of  $-t_r$ ,  $(r = 1, 2, \dots, p)$  such that  $w_r > 0$  and  $\oplus_{r=1}^p w_r = 1$  and  $(\sigma(1), \sigma(2), \dots, \sigma(p))$  are the permutations of  $-t_r$ ,  $(r = 1, 2, \dots, p)$  for which  $-t_{\sigma(r-1)} \geq -t_{\sigma(r)}$  for all  $r = 1, 2, \dots, p$ .

**Remark 7.** In Theorem 8,

- (i) If  $\delta_r = 0$ , for all  $r$ , then the (TFDOWG) operator reduces to the IFDOWG operator, and

$$\begin{aligned} & \text{IFDOWG}_w(-t_1, -t_2, \dots, -t_p) \\ &= \otimes_{r=1}^p (-t_{\sigma(r)})^{w_r} = \left( \begin{array}{l} \frac{1}{1 + \left\{ \oplus_{r=1}^p \Phi_r(1 - \mu_{\sigma(r)}/\mu_{\sigma(r)})^{\mathfrak{R}} \right\}^{(1/\mathfrak{R})}}, 1 - \frac{1}{1 + \left\{ \oplus_{r=1}^p \Phi_r(\lambda_{\sigma(r)}/1 - \lambda_{\sigma(r)})^{\mathfrak{R}} \right\}^{(1/\mathfrak{R})}} \end{array} \right). \end{aligned} \quad (41)$$

- (ii) If  $\lambda_r = 0$ , for all  $r$ , then the TFDOWG operator reduces to the BFDOWG operator, and

$$\begin{aligned} & \text{BFDOWG}_w(-t_1, -t_2, \dots, -t_p) \\ &= \otimes_{r=1}^p (-t_{\sigma(r)})^{w_r} = \left( \begin{array}{l} \frac{1}{1 + \left\{ \oplus_{r=1}^p \Phi_r(1 - \mu_{\sigma(r)}/\mu_{\sigma(r)})^{\mathfrak{R}} \right\}^{(1/\mathfrak{R})}}, -1 + \frac{1}{1 + \left\{ \oplus_{r=1}^p \Phi_r(|\delta_{\sigma(r)}|/1 + \delta_{\sigma(r)})^{\mathfrak{R}} \right\}^{(1/\mathfrak{R})}} \end{array} \right). \end{aligned} \quad (42)$$

The TFDOWG operators have the following properties.

(P1) (idempotency property). If  $-t_r = (\mu_r, \lambda_r, \delta_r)$ ,  $(r = 1, \dots, p)$ , is a family of TFNs such that they all equal, i.e., if  $-t_r = (\mu_r, \lambda_r, \delta_r) = -t = (\mu, \lambda, \delta)$  for all  $r$ , then

$$\text{TFDOWG}_\Phi(-t_1, -t_2, \dots, -t_r) = -t. \quad (43)$$

(P2) (boundedness property). Let  $-t_r = (\mu_r, \lambda_r, \delta_r)$ ,  $(r = 1, 2, \dots, p)$ , be a family of TFNs and  $-t^- = \min -t_r$  and  $-t^+ = \max -t_r$ . Then,

$$-t^- \leq \text{TFDOWG}_\Phi(-t_1, -t_2, \dots, -t_p) \leq -t^+. \quad (44)$$

(P3) (monotonicity property). If  $-t_r = (\mu_r, \lambda_r, \delta_r)$ ,  $(r = 1, 2, \dots, p)$ , and  $-t'_r = (\mu'_r, \lambda'_r, \delta'_r)$ ,

( $r = 1, 2, \dots, p$ ), are two TFNs such that  $-t_r \leq -t'_r$  for all  $r$ , then

$$\text{TFDOWG}_\Phi(-t_1, -t_2, \dots, -t_p) \leq \text{TFDOWG}_\Phi(-t'_1, -t'_2, \dots, -t'_p). \quad (45)$$

**Definition 15.** A TFDHWG operator of dimension  $p$  is a mapping  $\text{TFDHWG}: -t^p \rightarrow t$  with associated weight vector

$w = (w_1, w_2, \dots, w_p)$  such that  $w_r > 0$ ,  $\oplus_{r=1}^p w_r = 1$ , and TFDHWG operator can be evaluated as

$$\begin{aligned} & \text{TFDHWG}_{(w, \Phi)}(-t_1, -t_2, \dots, -t_p) \\ &= \otimes_{r=1}^p (-t_{\sigma(r)})^{w_r} = \left( \begin{array}{l} \frac{1}{1 + \left\{ \oplus_{r=1}^p w_r (1 - \mu_{\sigma(r)} / \mu_{\sigma(r)})^{\mathfrak{R}} \right\}^{(1/\mathfrak{R})}}, 1 - \frac{1}{1 + \left\{ \oplus_{r=1}^p w_r (\lambda_{\sigma(r)} / 1 - \lambda_{\sigma(r)})^{\mathfrak{R}} \right\}^{(1/\mathfrak{R})}}, \\ -1 + \frac{1}{1 + \left\{ \oplus_{r=1}^p w_r (|\delta_{\sigma(r)}| / 1 - \delta_{\sigma(r)})^{\mathfrak{R}} \right\}^{(1/\mathfrak{R})}} \end{array} \right), \end{aligned} \quad (46)$$

where  $-t_{\sigma(r)}$  is the  $r^{\text{th}}$  largest weighted tripolar fuzzy number, denoted by  $-t_r$ , and  $(-t_r = p\Phi - t_r, r = 1, 2, \dots, p)$  and  $\Phi = (\Phi_1, \Phi_2, \dots, \Phi_p)^T$  are the weight vectors of  $-t_r$  with  $\Phi_r > 0$ ,  $\oplus_{r=1}^p \Phi_r = 1$ , where  $p$  is the balancing coefficient. A TFDHWG operator has some special cases:

As a result, the TFDHG operator is a generalization of both the TFDWG and TFDOWG operators, reflecting the degrees of the provided arguments as well as their ordered locations.

**Remark 8.** In Definition 15,

- (i) When  $w = ((1/p), (1/p), \dots, (1/p))$ , then (TFDWG) operator is a special case of TFDHWG operator.
- (ii) When  $\Phi = ((1/p), (1/p), \dots, (1/p))$ , then (TFDOWG) operator becomes a special case of TFDHWG operator.

- (i) If  $\delta_r = 0$ , for all  $r = 1, 2, \dots, p$ , then TFDHWG operator reduces to IFDHWG operator and

$$\begin{aligned} & \text{IFDHWG}_{(w, \Phi)}(-t_1, -t_2, \dots, -t_p) \\ &= \otimes_{r=1}^p (-t_{\sigma(r)})^{w_r} = \left( \frac{1}{1 + \left\{ \oplus_{r=1}^p w_r (1 - \mu_{\sigma(r)} / \mu_{\sigma(r)})^{\mathfrak{R}} \right\}^{(1/\mathfrak{R})}}, 1 - \frac{1}{1 + \left\{ \oplus_{r=1}^p w_r (\lambda_{\sigma(r)} / 1 - \lambda_{\sigma(r)})^{\mathfrak{R}} \right\}^{(1/\mathfrak{R})}} \right). \end{aligned} \quad (47)$$

- (ii) If  $\lambda_r = 0$ , for all  $r$ , then TFDHWG operator reduces to BFDHWG operator and



$$\begin{aligned} & \text{BFDHWG}_{(w,\Phi)}(-t_1, -t_2, \dots, -t_p) \\ &= \otimes_{r=1}^p (-t_{\sigma(r)})^{w_r} = \left( \frac{1}{1 + \left\{ \oplus_{r=1}^p w_r (1 - \mu_{\sigma(r)}^{\delta_{\sigma(r)}} / \mu_{\sigma(r)}^{\delta_{\sigma(r)}}) \right\}^{\mathfrak{R}}} \right)^{(1/\mathfrak{R})}, -1 + \frac{1}{1 + \left\{ \oplus_{r=1}^p w_r (|\delta_{\sigma(r)}| / 1 - \delta_{\sigma(r)}) \right\}^{\mathfrak{R}}} \right)^{(1/\mathfrak{R})}. \end{aligned} \quad (48)$$

For the validity of Definition 15, we consider the following example.

*Example 2.* Let  $-t_1 = (0.5, 0.4, -0.3)$ ,  $-t_2 = (0.6, 0.3, -0.3)$ ,  $-t_3 = (0.7, 0.2, -0.3)$ , and  $-t_4 = (0.2, 0.3, -0.4)$  be four

TFNs and  $\Phi = (0.20, 0.30, 0.30, 0.20)^T$  and  $w = (0.2, 0.1, 0.3, 0.4)^T$  be the weight vector of TFNs and associated weight vector, respectively. Then, by Definition 15, for aggregated value of TFNs for  $(\mathfrak{R} = 3)$  and by using TFDHG operator, we have

$$\begin{aligned} -t_1 &= \left( \frac{1}{1 + \{4 \times 0.2 (1 - 0.5/0.5)^3\}^{(1/3)}}, 1 - \frac{1}{1 + \{4 \times 0.2 (0.4/1 - 0.4)^3\}^{(1/3)}}, \right. \\ &\quad \left. -1 + \frac{1}{1 + \{4 \times 0.20 (0.3/1 - 0.3)^3\}^{(1/3)}} \right) \\ &= (0.4814, 0.4179, -0.3158), \\ -t_2 &= \left( \frac{1}{1 + \{4 \times 0.3 (1 - 0.6/0.6)^3\}^{(1/3)}}, 1 - \frac{1}{1 + \{4 \times 0.3 (0.3/1 - 0.3)^3\}^{(1/3)}}, \right. \\ &\quad \left. -1 + \frac{1}{1 + \{4 \times 0.30 (0.3/1 - 0.3)^3\}^{(1/3)}} \right) \\ &= (0.5820, 0.3158, -0.3158), \\ -t_3 &= \left( \frac{1}{1 + \{4 \times 0.3 (1 - 0.7/0.7)^3\}^{(1/3)}}, 1 - \frac{1}{1 + \{4 \times 0.3 (0.2/1 - 0.2)^3\}^{(1/3)}}, \right. \\ &\quad \left. -1 + \frac{1}{1 + \{4 \times 0.30 (0.3/1 - 0.3)^3\}^{(1/3)}} \right) \\ &= (0.6841, 0.2121, -0.3158), \\ -t_4 &= \left( \frac{1}{1 + \{4 \times 0.2 (1 - 0.2/0.2)^3\}^{(1/3)}}, 1 - \frac{1}{1 + \{4 \times 0.2 (0.3/1 - 0.3)^3\}^{(1/3)}}, \right. \\ &\quad \left. -1 + \frac{1}{1 + \{4 \times 0.20 (0.4/1 - 0.4)^3\}^{(1/3)}} \right) \\ &= (0.1883, 0.3158, -0.4179). \end{aligned} \quad (49)$$

Scores of  $-t_r$ , ( $r = 1, 2, 3, 4$ ) can be approximated:

$$\begin{aligned}
 \textcircled{S}(-t'_1) &= \frac{1 + 0.4814 + 0.4179 - 0.3158}{3} = 0.5278, \\
 \textcircled{S}(-t'_2) &= \frac{1 + 0.5820 + 0.3158 - 0.3158}{3} = 0.5273, \\
 \textcircled{S}(-t'_3) &= \frac{1 + 0.6841 + 0.2121 - 0.3158}{3} = 0.5268, \\
 \textcircled{S}(-t'_4) &= \frac{1 + 0.1883 + 0.3851 - 0.4179}{3} = 0.3620.
 \end{aligned} \tag{50}$$

Since  $\textcircled{S}(-t'_1) > \textcircled{S}(-t'_2) > \textcircled{S}(-t'_3) > \textcircled{S}(-t'_4)$ , then we have

$$\begin{aligned}
 -t'_{\sigma(1)} &= -t'_1 = (0.4814, 0.4179, -0.3158), & -t'_{\sigma(2)} &= -t'_2 = (0.5820, 0.3158, -0.3158), \\
 -t'_{\sigma(3)} &= -t'_3 = (0.6841, 0.2121, -0.3158), & -t'_{\sigma(4)} &= -t'_4 = (0.1883, 0.3158, -0.4179).
 \end{aligned} \tag{51}$$

According to Definition 15, the aggregated values under TFDHWG operators for  $(R = 3)$  are calculated as

$$\begin{aligned}
 &\text{TFDHWG}_{(w,\Phi)}(-t_1, -t_2, \dots, -t_4) \\
 &= \otimes_{r=1}^4 (-t'_{\sigma(r)})^{w_r} = \left( \begin{array}{c} \frac{1}{1 + \left\{ \sum_{r=1}^4 w_r (1 - \mu'_{\sigma(r)}/\mu'_{\sigma(r)})^3 \right\}^{(1/3)}}, 1 - \frac{1}{1 + \left\{ \sum_{r=1}^4 (\lambda'_{\sigma(r)}/1 - \lambda'_{\sigma(r)})^3 \right\}^{(1/3)}}, \\ -1 + \frac{1}{1 + \left\{ \sum_{r=1}^4 (|\delta'_{\sigma(r)}|/1 + \delta'_{\sigma(r)})^3 \right\}^{(1/3)}} \end{array} \right) \\
 &= \left( \begin{array}{c} \frac{1}{1 + \left\{ 0.2(1 - 0.4814/0.4814)^3 + 0.1(1 - 0.5820/0.5820)^3 + 0.3(1 - 0.6841/0.6841)^3 + 0.4(1 - 0.1883/0.1883)^3 \right\}^{(1/3)}}, \\ 1 - \frac{1}{1 + \left\{ 0.2(0.4179/1 - 0.4179)^3 + 0.1(0.3158/1 - 0.3158)^3 + 0.3(0.2121/1 - 0.2121)^3 + 0.4(0.3158/1 - 0.3158)^3 \right\}^{(1/3)}}, \\ -1 + \frac{1}{1 + \left\{ 0.2(0.3158/1 - 0.3158)^3 + 0.1(0.3158/1 - 0.3158)^3 + 0.3(0.3158/1 - 0.3158)^3 + 0.4(0.4179/1 - 0.4179)^3 \right\}^{(1/3)}} \end{array} \right) \\
 &= (0.0799, 0.2351, -0.3452).
 \end{aligned} \tag{52}$$

### 3. Models for MADM with TF Information

To solve a MADM problem using TF information, we will use the TFDA operators established in the preceding sections. The MADM problem is represented using the following technique or notations for prospective evaluation of developing technology commercialization using TF information. Let  $A = \{A_1, A_2, \dots, A_m\}$  present a discrete set of alternatives and  $G = \{G_1, G_2, \dots, G_n\}$ , the set of attributes. Let  $\Phi = (\Phi_1, \Phi_2, \dots, \Phi_n)$  be the weight of attributes in the form of real numbers with the conditions  $\Phi_j > 0$  and  $\sum_{j=1}^n \Phi_j = 1$ . Suppose that  $M = (\tilde{m}_{ij})_{m \times n} = (\mu_{ij}, \lambda_{ij}, \delta_{ij})_{m \times n}$  is the TF decision matrix, where  $\mu_{ij}$  represent the degree of

membership,  $\lambda_{ij}$  represent the degree of non-membership, and  $\delta_{ij}$  denote the degree of negative membership such that,  $\mu_{ij} + \lambda_{ij} \leq 1$  and  $\mu_{ij}, \lambda_{ij} \in [0, 1]$  and  $\delta_{ij} \in [-1, 0]$ , ( $i = 1, 2, \dots, n$ ) and ( $j = 1, 2, \dots, m$ ). The process of utilizing the TFDWA (or TFDWG) operator to solve an MADM problem is listed below (Algorithm 1).

### 4. Illustrative Example

To present the practical output of the constructed model, we consider the MADM problem (adapted from [27]) of the ERP systems. We suppose that an organization wants to employ the ERP systems. The first step for the ERP

*Step 1.* We apply the TFDWA operator to process the information in the decision matrix  $M$  and to compute the overall values of the alternatives  $A_i$  ( $i = 1, 2, \dots, m$ ), and we have

$$\begin{aligned} \tilde{m}_i = (\mu_i, \lambda_i, \delta_i) &= \text{TFDWA}_\Phi(\tilde{m}_{i1}, \tilde{m}_{i2}, \dots, \tilde{m}_{in}) = \Phi_{j=1}^n (\Phi_j \tilde{m}_{ij}) \\ &= \left( \begin{array}{l} 1 - (1/1 + \left\{ \sum_{j=1}^n \Phi_j (\mu_{ij}/1 - \mu_{ij})^{\mathfrak{R}} \right\}^{(1/\mathfrak{R})}), (1/1 + \left\{ \sum_{j=1}^n \Phi_j (1 - \lambda_{ij}/\lambda_{ij})^{\mathfrak{R}} \right\}^{(1/\mathfrak{R})}), \\ (-1/1 + \left\{ \sum_{j=1}^n \Phi_j (1 + \delta_{ij}/|\delta_{ij}|)^{\mathfrak{R}} \right\}^{(1/\mathfrak{R})}). \end{array} \right) \end{aligned}$$

Or if we select the TFDWG operator, instead of TFDWA operator, then we have

$$\begin{aligned} \tilde{m}_i = (\mu_i, \lambda_i, \delta_i) &= \text{TFDWG}_\Phi(\tilde{m}_{i1}, \tilde{m}_{i2}, \dots, \tilde{m}_{in}) = \otimes_{j=1}^n (\tilde{m}_{ij})^{\Phi_j} \\ &= \left( \begin{array}{l} (1/1 + \left\{ \sum_{j=1}^n \Phi_j (1 - \mu_{ij}/\mu_{ij})^{\mathfrak{R}} \right\}^{(1/\mathfrak{R})}), 1 - (1/1 + \left\{ \sum_{j=1}^n \Phi_j (\lambda_{ij}/1 - \lambda_{ij})^{\mathfrak{R}} \right\}^{(1/\mathfrak{R})}), \\ -1 + (1/1 + \left\{ \sum_{j=1}^n \Phi_j (|\delta_{ij}|/1 + \delta_{ij})^{\mathfrak{R}} \right\}^{(1/\mathfrak{R})}). \end{array} \right) \end{aligned}$$

*Step 2.* Find the score values  $\odot(\tilde{m}_i)$  ( $i = 1, 2, \dots, m$ ) of all the alternatives.

*Step 3.* Rank all  $A_i$  ( $i = 1, 2, \dots, m$ ) according to the score values  $\odot(\tilde{m}_i)$ , ( $i = 1, 2, \dots, m$ ). In case  $\odot(\tilde{m}_i)$  and  $\odot(\tilde{m}_j)$  values are equal, then calculate the accuracy values  $\text{acc}(\tilde{m}_i)$  and  $\text{acc}(\tilde{m}_j)$  to rank the alternatives  $A_i$  and  $A_j$ , respectively.

*Step 4.* Select the best alternative(s).

#### ALGORITHM 1: MADM steps.

system is to construct a team of experts which contained CIO and two other senior experts from the source institute. The expert's team collects all the data regarding ERP vendors and system, in the form of TFS. Secondly, the team selects five potential ERP systems, treated as  $A_i$ , ( $i = 1, 2, 3, 4, 5$ ), as candidates. The team also imposes four attributes for the evaluation of this MADM problem of  $A_i$  ( $i = 1, 2, 3, 4, 5$ ):

- $G_1$  = function and technology.
- $G_2$  = strategic fitness.
- $G_3$  = vendor's reputation.
- $G_4$  = vendor's ability.

The expert team gave some weighting for the selection of optimal candidate as  $\Phi = (0.2, 0.1, 0.3, 0.4)$ , and the information collected is in the form of TFSs, satisfying the above four attributes. The rating of the team is presented in the matrix given below.

To find the most favorable ERP systems, we employ the TFDWA operator (or TFDWG operator) in the following to construct a solution to MADM problems based on TF information, which may be summarized as follows.

*Step 1.* Let  $\mathfrak{R} = 1$ . By applying the TFDWA operator, we compute the overall preference values  $\tilde{m}_i$  of the ERP system  $A_i$  ( $i = 1, 2, 3, 4, 5$ ):

$$\begin{aligned} \tilde{m}_1 &= (0.0013, 0.9455, -0.9871), \\ \tilde{m}_2 &= (0.0010, 0.9800, -0.9964), \\ \tilde{m}_3 &= (0.0004, 0.9788, -0.8852), \\ \tilde{m}_4 &= (0.0001, 0.9976, -0.9936), \\ \tilde{m}_5 &= (0.0037, 0.9562, -0.9814). \end{aligned} \tag{53}$$

*Step 2.* Find the score values  $\odot(\tilde{m}_i)$ , ( $i = 1, 2, 3, 4, 5$ ), of the overall TF numbers, TFNs,  $\tilde{m}_i$ , ( $i = 1, 2, 3, 4, 5$ ):

$$\begin{aligned} \odot(\tilde{m}_1) &= 0.3199, \\ \odot(\tilde{m}_2) &= 0.3282, \\ \odot(\tilde{m}_3) &= 0.3646, \\ \odot(\tilde{m}_4) &= 0.3347, \\ \odot(\tilde{m}_5) &= 0.3261. \end{aligned} \tag{54}$$

*Step 3.* Rank all the ERP systems  $A_i$ , ( $i = 1, 2, 3, 4, 5$ ), according to the respective score values of the overall tripolar fuzzy numbers, and we get  $A_3 > A_4 > A_2 > A_5 > A_1$ .

*Step 4.* According to score values,  $A_3$  is the most considerable ERP system.

If we use the TFDWG operator instead of TFDWA operator, then we solve the problem similarly as above:

*Step 1.* For  $\mathfrak{R} = 1$ , aggregate all TFNs via the TFDWG operator to derive the overall TFNs  $\tilde{m}_i$  ( $i = 1, 2, 3, 4, 5$ ) of the ERP system:

$$\begin{aligned}\tilde{m}_1 &= (0.9959, 0.0000, -0.0004), \\ \tilde{m}_2 &= (0.9944, 0.0002, -0.0015), \\ \tilde{m}_3 &= (0.9875, 0.0002, -0.0000), \\ \tilde{m}_4 &= (0.9674, 0.0023, -0.0008), \\ \tilde{m}_5 &= (0.9984, 0.0001, -0.0003).\end{aligned}\quad (55)$$

*Step 2.* Calculate the score values  $\textcircled{\text{S}}(\tilde{m}_i)$ , ( $i = 1, 2, 3, 4, 5$ ), of the overall TF numbers, TFNs,  $\tilde{m}_i$ , ( $i = 1, 2, 3, 4, 5$ ):

$$\begin{aligned}\textcircled{\text{S}}(\tilde{m}_1) &= 0.6651, \\ \textcircled{\text{S}}(\tilde{m}_2) &= 0.6643, \\ \textcircled{\text{S}}(\tilde{m}_3) &= 0.6625, \\ \textcircled{\text{S}}(\tilde{m}_4) &= 0.6563, \\ \textcircled{\text{S}}(\tilde{m}_5) &= 0.6660.\end{aligned}\quad (56)$$

*Step 3.* Rank all the ERP systems  $A_i$ , ( $i = 1, 2, 3, 4, 5$ ), according to the respective score values of the overall TF numbers, and we get  $\textcircled{\text{S}}A_5 > \textcircled{\text{S}}A_1 > \textcircled{\text{S}}A_2 > \textcircled{\text{S}}A_3 > \textcircled{\text{S}}A_4$ .

*Step 4.* By the score values,  $A_5$  is the most ideal ERP system.

Based on the above discussion, it is determined that, while ranking values of the alternatives change when utilizing the TFDWA (TFDWG) operator, the ranking orders of the ERP system did not remain the same, i.e., the most ideal ERP system through using TFDWA operator is  $A_4$ , while using TFDWG operator, it is  $A_5$ . The reason for different optimal alternatives is the use of different methods in the decision process.

In Figure 1, the comparison of the ERP system is shown, where it is observed that the graph of TFDWA operator (green line) has some fluctuations in the middle values of ranking order of ERP system while the graph of TFDWG operator (brown line) is stable. This means that the ranking order of ERP systems is more stable in TFDWG operator as compared to the ranking order of ERP systems in TFDWA operator.

**4.1. Influence of Parameter.** To describe the impact of the operational parameters  $\mathfrak{R}$  on MADM outcomes, we will rank the alternatives using different values of  $\mathfrak{R}$ . The results of score function and ranking order of the ERP system  $A_i$  ( $i = 1, 2, 3, 4, 5$ ) in the range of  $\mathfrak{R} \in [1, 10]$  applying TFDWA and TFDWG operators are presented. The corresponding scores and ranking of the ERP system  $A_i$ , ( $i = 1, 2, 3, 4, 5$ ), are shown in Tables 1 and 2.

Table 3 shows that altering the value of  $\mathfrak{R}$  in the TFDWA operator changes the ranking orders, and the associated best alternatives are not similar. Several ranking orders of the

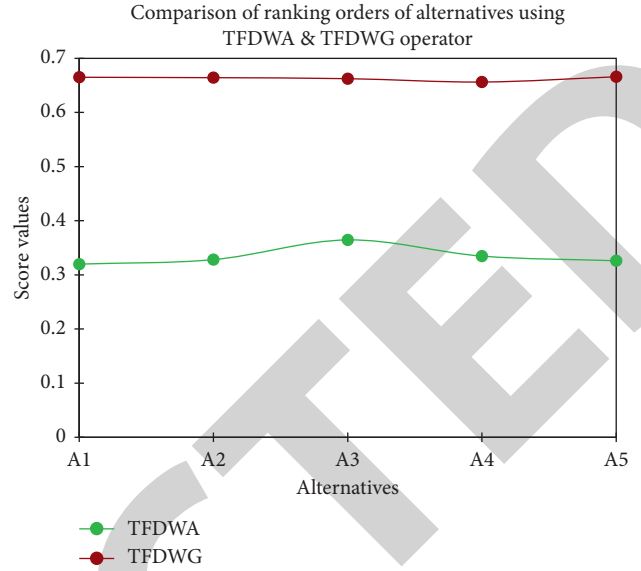


FIGURE 1: Graphs of TFDWA and TFDWG operators.

TABLE 1: Tripolar fuzzy (TF) decision matrix with its values.

TF matrix values				
$M =$	$(0.7, 0.2, -0.3)$	$(0.2, 0.4, -0.8)$	$(0.2, 0.5, -0.3)$	$(0.8, 0.2, -0.2)$
	$(0.6, 0.3, -0.4)$	$(0.3, 0.6, -0.5)$	$(0.5, 0.3, -0.4)$	$(0.4, 0.3, -0.6)$
	$(0.4, 0.5, -0.2)$	$(0.4, 0.4, -0.1)$	$(0.3, 0.2, -0.5)$	$(0.5, 0.4, -0.4)$
	$(0.2, 0.6, -0.6)$	$(0.3, 0.4, -0.7)$	$(0.4, 0.6, -0.2)$	$(0.5, 0.4, -0.3)$
	$(0.7, 0.3, -0.3)$	$(0.5, 0.4, -0.4)$	$(0.5, 0.3, -0.4)$	$(0.4, 0.3, -0.4)$

alternatives for subintervals of  $[2, 30]$  by using TFDWA and TFDWG operators of the operational parameter  $\mathfrak{R}$  are shown in Tables 2 and 4.

Ranking orders of the alternatives for subintervals of  $[2, 30]$  by using TFDWG operators of the operational parameter  $\mathfrak{R}$  are shown in Table 5.

The corresponding graphs in Figures 2 and 3 are provided to diagnose the influence of parameter  $\mathfrak{R} \in [1, 10]$  on the ranking of alternatives in TFDWA and TFDWG operators.

From Figures 2 and 3, we observe that the ranking order of the ERP systems in TFDWA operator (Figure 2) has more fluctuations than the ranking order of alternatives in TFDWG operators (Figure 3). Therefore, the proposed method of TFDWG operator produced stable results of the ERP system.

**4.2. Comparative Analysis.** To compare the developed models with the existing literature to describe the advantages of the proposed models, we consider the following two special cases of the suggested models, given in Examples 1 and 2. In Example 1, we dropped the implicit counter property from the information given in Table 1 and the reduced matrix becomes the IF information which is given in Table 6.

TABLE 2: Ranking order in TFDWA operator.

$\mathfrak{R}$	Ranking orders
$\mathfrak{R} \in 1, 2]$	$\otimes A_3 > \otimes A_4 > \otimes A_2 > \otimes A_5 > \otimes A_1$
$\mathfrak{R} \in (2, 4]$	$\otimes A_3 > \otimes A_4 > \otimes A_1 > \otimes A_5 > \otimes A_2$
$\mathfrak{R} \in (4, 5]$	$\otimes A_5 > \otimes A_3 > \otimes A_4 > \otimes A_1 > \otimes A_2$
$\mathfrak{R} \in (5, 6]$	$\otimes A_4 > \otimes A_3 > \otimes A_5 > \otimes A_1 > \otimes A_2$
$\mathfrak{R} \in (6, 10]$	$\otimes A_5 > \otimes A_4 > \otimes A_3 > \otimes A_1 > \otimes A_2$

TABLE 3: The ERP's ranking order for various working parameters in TFDWA operator.

$\mathfrak{R}$	$\otimes (\tilde{m}_1), \otimes (\tilde{m}_2), \otimes (\tilde{m}_3), \otimes (\tilde{m}_4), \otimes (\tilde{m}_5)$	Ranking order
1	0.3199, 0.3282, 0.3646, 0.3347, 0.3261	$\otimes A_3 > \otimes A_4 > \otimes A_2 > \otimes A_5 > \otimes A_1$
2	0.2326, 0.2652, 0.4763, 0.3574, 0.2887	$\otimes A_3 > \otimes A_4 > \otimes A_5 > \otimes A_2 > \otimes A_1$
3	0.2438, 0.2300, 0.4522, 0.3848, 0.2335	$\otimes A_3 > \otimes A_4 > \otimes A_1 > \otimes A_5 > \otimes A_2$
4	0.2730, 0.2279, 0.4324, 0.4018, 0.2091	$\otimes A_3 > \otimes A_4 > \otimes A_1 > \otimes A_5 > \otimes A_2$
5	0.2968, 0.2354, 0.4221, 0.4114, 0.4257	$\otimes A_5 > \otimes A_3 > \otimes A_4 > \otimes A_1 > \otimes A_2$
6	0.3146, 0.2446, 0.4166, 0.4171, 0.3896	$\otimes A_4 > \otimes A_3 > \otimes A_5 > \otimes A_1 > \otimes A_2$
7	0.3281, 0.2531, 0.4135, 0.4206, 0.4253	$\otimes A_5 > \otimes A_4 > \otimes A_3 > \otimes A_1 > \otimes A_2$
8	0.3386, 0.2607, 0.4118, 0.4230, 0.4367	$\otimes A_5 > \otimes A_4 > \otimes A_3 > \otimes A_1 > \otimes A_2$
9	0.3470, 0.2671, 0.4107, 0.4246, 0.4457	$\otimes A_5 > \otimes A_4 > \otimes A_3 > \otimes A_1 > \otimes A_2$
10	0.3537, 0.2727, 0.4101, 0.4257, 0.4529	$\otimes A_5 > \otimes A_4 > \otimes A_3 > \otimes A_1 > \otimes A_2$

TABLE 4: Ranking order in TFDWG operator.

$\mathfrak{R}$	Ranking order
$\mathfrak{R} \in 0, 3]$	$\otimes A_5 > \otimes A_1 > \otimes A_2 > \otimes A_3 > \otimes A_4$
$\mathfrak{R} \in (3, 5]$	$\otimes A_3 > \otimes A_4 > \otimes A_1 > \otimes A_5 > \otimes A_2$
$\mathfrak{R} \in (5, 7]$	$\otimes A_5 > \otimes A_1 > \otimes A_2 > \otimes A_3 > \otimes A_4$
$\mathfrak{R} \in (7, 8]$	$\otimes A_5 > \otimes A_1 > \otimes A_3 > \otimes A_4 > \otimes A_2$
$\mathfrak{R} \in (8, 10]$	$\otimes A_5 > \otimes A_1 > \otimes A_3 > \otimes A_4 > \otimes A_2$

TABLE 5: Ranking order for various working parameters in TFDWG operator.

$\mathfrak{R}$	$\otimes (\tilde{m}_1), \otimes (\tilde{m}_2), \otimes (\tilde{m}_3), \otimes (\tilde{m}_4), \otimes (\tilde{m}_5)$	Ranking order
1	0.6651, 0.6643, 0.6625, 0.6563, 0.6660	$\otimes A_5 > \otimes A_1 > \otimes A_2 > \otimes A_3 > \otimes A_4$
2	0.6385, 0.6238, 0.5998, 0.5406, 0.6552	$\otimes A_5 > \otimes A_1 > \otimes A_2 > \otimes A_3 > \otimes A_4$
3	0.5982, 0.5651, 0.5331, 0.4726, 0.6369	$\otimes A_5 > \otimes A_1 > \otimes A_2 > \otimes A_3 > \otimes A_4$
4	0.5650, 0.5187, 0.4941, 0.4495, 0.6198	$\otimes A_3 > \otimes A_4 > \otimes A_1 > \otimes A_5 > \otimes A_2$
5	0.5403, 0.4855, 0.4718, 0.4495, 0.6195	$\otimes A_3 > \otimes A_4 > \otimes A_1 > \otimes A_5 > \otimes A_2$
6	0.5222, 0.4615, 0.4492, 0.4349, 0.5857	$\otimes A_5 > \otimes A_1 > \otimes A_2 > \otimes A_3 > \otimes A_4$
7	0.5083, 0.4438, 0.4492, 0.4349, 0.5857	$\otimes A_5 > \otimes A_1 > \otimes A_2 > \otimes A_3 > \otimes A_4$
8	0.4977, 0.4302, 0.4430, 0.4360, 0.5784	$\otimes A_5 > \otimes A_1 > \otimes A_3 > \otimes A_4 > \otimes A_2$
9	0.4892, 0.4194, 0.4385, 0.4333, 0.5725	$\otimes A_5 > \otimes A_1 > \otimes A_3 > \otimes A_4 > \otimes A_2$
10	0.4823, 0.4108, 0.4351, 0.4329, 0.5676	$\otimes A_5 > \otimes A_1 > \otimes A_3 > \otimes A_4 > \otimes A_2$

*Example 3.* Recall the TF information decision matrix given in Table 1. The modified decision matrix contained IF information given in Table 6, and the normalized decision matrix is shown in Table 7, respectively.

By applying IFDWA and IFDWG operators on Table 7, the aggregated IFNs are summarized in Table 8.

By calculating the score values of the aggregated IFNs of Table 9, we get the following table.

In Table 10, a comparison of the ranking of alternatives by using IFDWA operator and TFDWA operator is shown, and the graphical view of these operators is highlighted in Figure 4.

From Figure 4, we can see that the ranking order of alternatives for the TFDWA operator (brown line) is stable from  $A_1$  to  $A_2$ , increasing from  $A_2$  to  $A_3$ , decreasing from  $A_3$  to  $A_4$ , and stable from  $A_4$  to  $A_5$ . From  $A_4$  it becomes stable. On the other hand, if we consider the IFDWA operator (green line), the ranking order of alternatives is decreasing continuously from  $A_1$  to  $A_4$  and increasing from  $A_4$  to  $A_5$ , and hence we observe no stability in ranking produced by IFDWA operator. Therefore, the proposed method of TFDWA operator gives more stable ranking than the existing methods in literature.

A comparison of the ranking of alternatives by using IFDWG operator and TFDWG operator is shown in

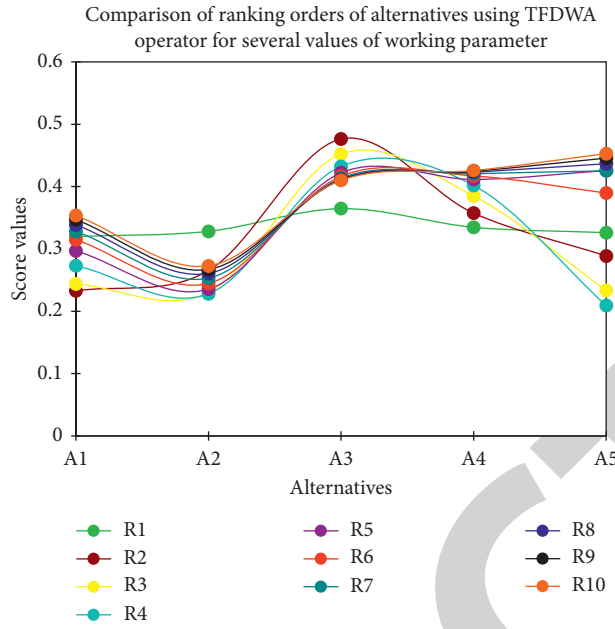


FIGURE 2: Graphs of alternatives in TFDWA operator ( $\alpha \in [2, 30]$ ).

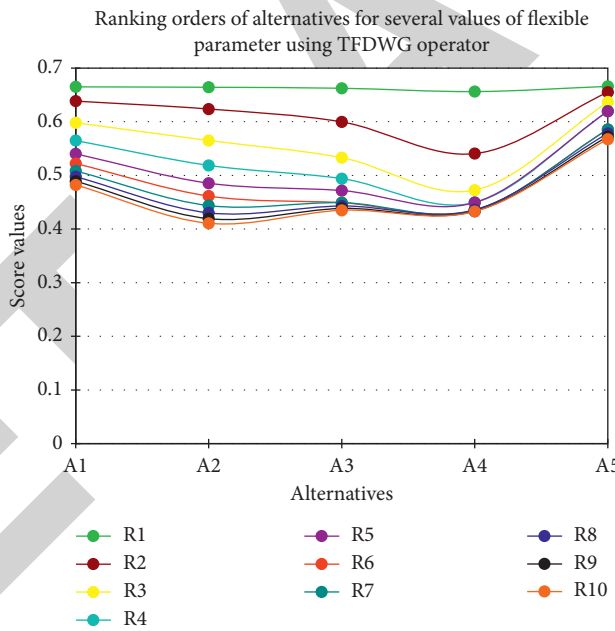


FIGURE 3: Graphs of alternatives in TFDWG operator ( $\alpha \in [2, 30]$ ).

TABLE 6: Intuitionistic fuzzy (IF) decision matrix.

IF matrix values

$$M' = \begin{bmatrix} (0.7, 0.2) & (0.2, 0.4) & (0.2, 0.5) & (0.8, 0.2) \\ (0.6, 0.3) & (0.3, 0.6) & (0.5, 0.3) & (0.4, 0.3) \\ (0.4, 0.5) & (0.4, 0.4) & (0.3, 0.2) & (0.5, 0.4) \\ (0.2, 0.6) & (0.3, 0.4) & (0.4, 0.6) & (0.5, 0.4) \\ (0.7, 0.3) & (0.5, 0.4) & (0.5, 0.3) & (0.4, 0.3) \end{bmatrix}$$

TABLE 7: Normalized intuitionistic fuzzy (NIF) decision matrix.

NIF matrix values

$$M = \begin{bmatrix} (0.7, 0.2) & (0.4, 0.2) & (0.2, 0.5) & (0.8, 0.2) \\ (0.6, 0.3) & (0.6, 0.3) & (0.5, 0.3) & (0.4, 0.3) \\ (0.4, 0.5) & (0.4, 0.4) & (0.3, 0.2) & (0.5, 0.4) \\ (0.2, 0.6) & (0.4, 0.3) & (0.4, 0.6) & (0.5, 0.4) \\ (0.7, 0.3) & (0.4, 0.5) & (0.5, 0.3) & (0.4, 0.3) \end{bmatrix}$$



TABLE 8: Aggregated IFNs using IFDWA and IFDWG operators.

Alternatives	BFDWA operator	BFDWG operator
A <sub>1</sub>	(0.0037, 0.8668)	(0.9984, 0.0000)
A <sub>2</sub>	(0.0030, 0.9335)	(0.9964, 0.0000)
A <sub>3</sub>	(0.0004, 0.9788)	(0.9875, 0.0002)
A <sub>4</sub>	(0.0002, 0.9962)	(0.9788, 0.0015)
A <sub>5</sub>	(0.0024, 0.9704)	(0.9976, 0.0001)

TABLE 9: Score values of IFNs using IFDWA and IFDWG operators.

Alternatives	IFDWA operator	IFDWG operator
A <sub>1</sub>	0.0684	0.9992
A <sub>2</sub>	0.0347	0.9982
A <sub>3</sub>	0.0108	0.9936
A <sub>4</sub>	0.0020	0.9886
A <sub>5</sub>	0.0160	0.9987

*Ranking order of alternatives*

IFDWA operator	$\textcircled{S}A_1 > \textcircled{S}A_2 > \textcircled{S}A_5 > \textcircled{S}A_3 > \textcircled{S}A_4$
IFDWG operator	$\textcircled{S}A_1 > \textcircled{S}A_5 > \textcircled{S}A_2 > \textcircled{S}A_3 > \textcircled{S}A_4$

TABLE 10: Ranking of alternatives using IFDWA and TFDWA operators.

IFDWA operator	$\textcircled{S}A_1 > \textcircled{S}A_2 > \textcircled{S}A_5 > \textcircled{S}A_3 > \textcircled{S}A_4$
TFDWA operator	$\textcircled{S}A_3 > \textcircled{S}A_4 > \textcircled{S}A_2 > \textcircled{S}A_5 > \textcircled{S}A_1$

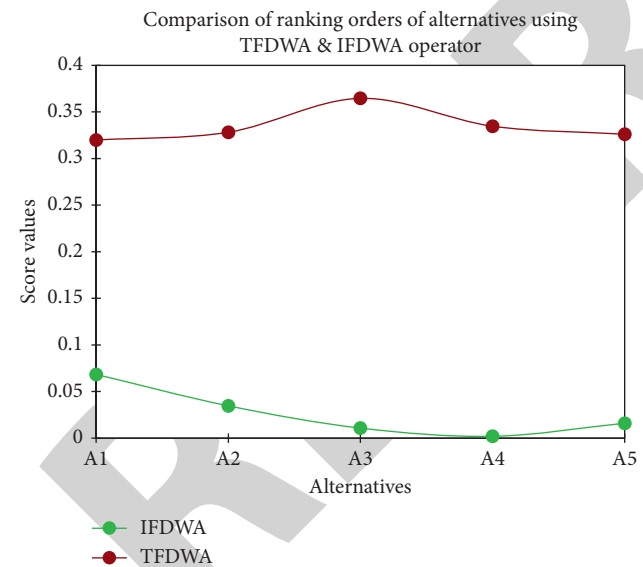


FIGURE 4: Comparison of ranking order in TFDWA and IFDWA operators.

Table 11, and the graphical view of these operators is highlighted in Figure 5.

From Figure 5, it is clearly seen that the ranking of alternatives in IFDWG operator (green line) is stable from A<sub>1</sub> to A<sub>5</sub>. Similarly, the ranking in TFDWG operator (brown line) is stable from A<sub>1</sub> to A<sub>5</sub>. Therefore, the stability of our proposed operator is similar to the stability of existing IFDWG operator, but the proposed TFDWG operator is more advanced than the existing operator, as it is a

TABLE 11: Ranking of alternatives using IFDWG and TFDWG operators.

IFDWG operator	$\textcircled{S}A_1 > \textcircled{S}A_2 > \textcircled{S}A_5 > \textcircled{S}A_3 > \textcircled{S}A_4$
TFDWG operator	$\textcircled{S}A_3 > \textcircled{S}A_4 > \textcircled{S}A_2 > \textcircled{S}A_5 > \textcircled{S}A_1$

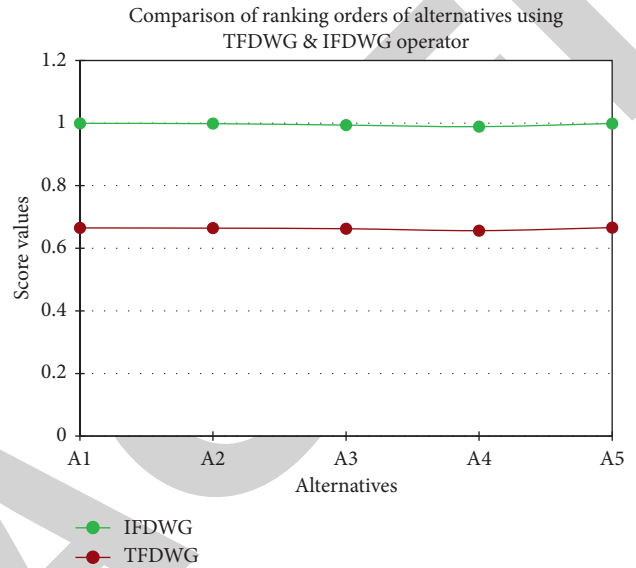


FIGURE 5: Ranking order of alternatives in TFDWG and IFDWG operators.

TABLE 12: Bipolar fuzzy (BF) decision matrix.

BF matrix values				
$M'' = \begin{bmatrix} (0.7, -0.3) & (0.2, -0.8) & (0.2, -0.3) & (0.8, -0.2) \\ (0.6, -0.4) & (0.3, -0.5) & (0.5, -0.4) & (0.4, -0.6) \\ (0.4, -0.2) & (0.4, -0.1) & (0.3, -0.5) & (0.5, -0.4) \\ (0.2, -0.6) & (0.3, -0.7) & (0.4, -0.2) & (0.5, -0.3) \\ (0.7, -0.3) & (0.5, -0.4) & (0.5, -0.4) & (0.4, -0.4) \end{bmatrix}$				

TABLE 13: Aggregated BFNs by using BFDWA and BFDWG operators.

Alternatives	BFDWA operator	BFDWG operator
A <sub>1</sub>	(0.0013, -0.9871)	(0.9875, -0.0004)
A <sub>2</sub>	(0.9959, -0.0004)	(0.0001, -0.9936)
A <sub>3</sub>	(0.0010, -0.9964)	(0.9674, -0.0008)
A <sub>4</sub>	(0.9944, -0.0015)	(0.0037, -0.9814)
A <sub>5</sub>	(0.0004, -0.8852)	(0.9984, -0.0003)

TABLE 14: Score values by using BFDWA and BFDWG operators.

Alternatives	BFDWA operator	BFDWG operator
A <sub>1</sub>	0.0071	0.9974
A <sub>2</sub>	0.0023	0.9964
A <sub>3</sub>	0.0576	0.9935
A <sub>4</sub>	0.0032	0.9833
A <sub>5</sub>	0.0111	0.9990

TABLE 15: Ranking of alternatives in BFDWA/BFDWG operators.

BFDWA operator	$\otimes A_3 > \otimes A_5 > \otimes A_1 > \otimes A_4 > \otimes A_2$
BFDWG operator	$\otimes A_5 > \otimes A_1 > \otimes A_2 > \otimes A_3 > \otimes A_4$

TABLE 16: Ranking of proposed and BFDWA/BFDWG operators.

BFDWA operator	$\otimes A_3 > \otimes A_5 > \otimes A_1 > \otimes A_4 > \otimes A_2$
BFDWG operator	$\otimes A_5 > \otimes A_1 > \otimes A_2 > \otimes A_3 > \otimes A_4$
TFDWA operator (proposed)	$\otimes A_3 > \otimes A_4 > \otimes A_2 > \otimes A_6 > \otimes A_1$
TFDWG operator (proposed)	$\otimes A_3 > \otimes A_4 > \otimes A_2 > \otimes A_5 > \otimes A_1$

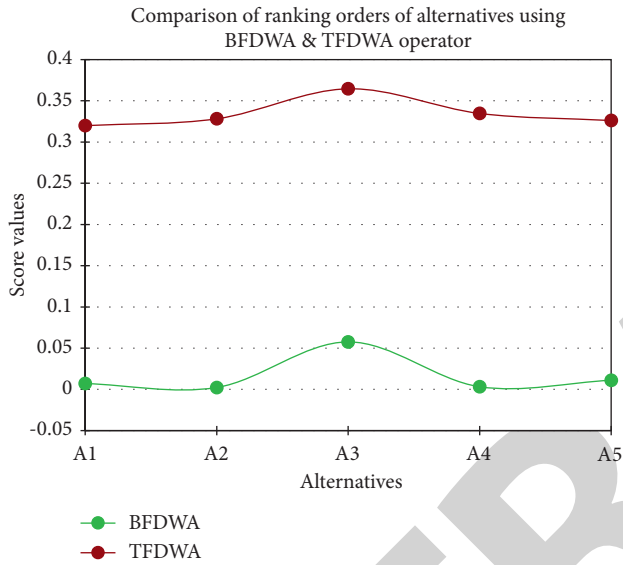


FIGURE 6: Comparison of alternatives in TFDWA and BFDWA operators.

generalization of both the notions of IFNs and BFNs and can handle problems involving IFN and BFN in the single model.

Now, we discuss the case of proposed operators by dropping the degree of non-membership from the information given in the decision matrix of Table 1. The following example illustrate the case of dropping the degree of non-membership.

*Example 4.* Reconsider the decision matrix given in Table 1 and drop the non-membership degree from each triplet of this matrix. The modified matrix is reduced to a BF decision matrix given in Table 12.

For the comparison of the proposed models, we apply BFDWA and BFDWG operators on Table 12 and the aggregated BFNs are summarized in Table 13.

The score values of the aggregated BFNs are summarized in Table 14.

The ranking of alternatives by using BFDWA/BFDWG operator is shown in Table 15.

In Table 16, we compare the ranking order of alternatives obtained by the proposed method and BFDWA/BFDWG operators.

In Figures 6 and 7, we compared the proposed operators and existing operators of BFDWA/BFDWG operators.

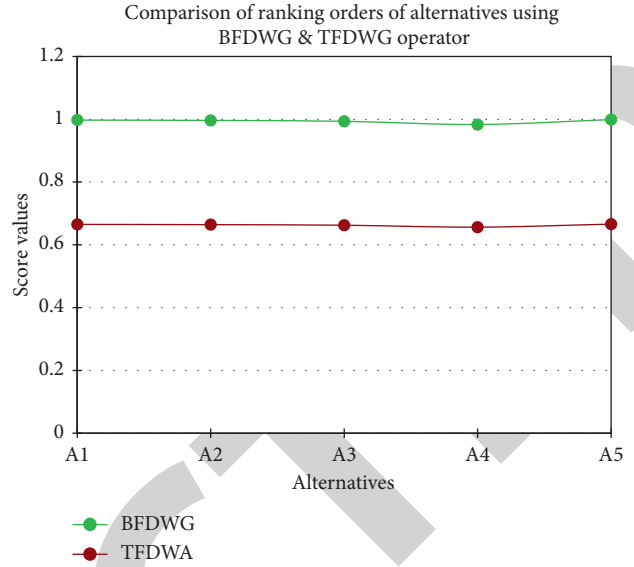


FIGURE 7: Comparison of alternatives in BFDWG and TFDWG operators.

From Figure 7, it is clearly seen that the ranking of alternatives using BFDWA operator (brown line) and TFDWA operator (green line) is not fully stable in between  $A_1$  to  $A_5$ .

From Figure 7, we observed that the ranking order of alternatives in BFDWG operator (green line) and TFDWG operator (brown line) is stable. This means that the proposed operators and existing operator behave similar. But the proposed operators are generalization of both the IFS and BFS, which combine both the notions into a single model.

### 5. Conclusion

In this article, we addressed MADM problems with TF information. From the inspiration of Dombi operations, we incorporated arithmetic and geometric operations to develop certain TF Dombi aggregation operators:

- (i) TFDWA operator.
- (ii) TFDOWA operator.
- (iii) TFDHWA operator.
- (iv) PFDWG operator.
- (v) TFDOWG operator.
- (vi) TFDHWG operator.

Different aspects of these suggested operators are highlighted. Then, we utilized these operators to broaden a few mechanisms to cope with MADM problems. Ultimately, a realistic example of an ERP system is provided to develop the strategy and show the effectiveness of the proposed method. In future research, we will investigate DM theory, risk theory, and other areas in an ambiguous environment for the suggested tripolarity models.

The method introduced in this paper is a generalization of IF and BF environments, and the present model deals with the tripolar fuzzy environment. The existing methods of IF

TABLE 17: Comparison of proposed and existing methods.

Methods	Describe fuzzy information easier	More flexible by a parameter
Dombi [30]	Yes	No
Xu and Yager [5]	Yes	No
Proposed operator	Yes	Yes

sets and BF sets [5, 30] for MADM problems have restrictions; in IFS, we cannot discuss the implicit counter property of objects while the BFS models failed to consider the non-membership property of objects in decision process, and definitely these deficiencies in IF models as well as in BF models will lose the information in the process. The present method combines the two concepts of IF and BF models together. In a nutshell, the suggested MADM technique for TFDWA and TFDWGA operators has improved DM reliability. Our suggested approaches are more comprehensive and adaptable than conventional methods for controlling IF and BF models. We compared our method with Dombi's method [30] and Xu and Yager's method [5]. The comparative results are shown in Table 17. As a conclusion, our suggested approaches are more flexible and adaptable than available methods for controlling IF and BF MADM challenges.

### Data Availability

The data used in the manuscript are hypothetical and can be used by anyone by just citing this article.

### Conflicts of Interest

The authors declare that they have no conflicts of interest.

### Acknowledgments

This research was supported by the Researchers Supporting Project number (RSP-2021/244), King Saud University, Riyadh, Saudi Arabia.

### References

- [1] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, no. 3, pp. 338–353, 1965.
- [2] K. T. Atanassov, "Intuitionistic fuzzy sets," *Fuzzy Sets and Systems*, vol. 20, no. 1, pp. 87–96, 1986.
- [3] K. T. Atanassov, "More on intuitionistic fuzzy sets," *Fuzzy Sets and Systems*, vol. 33, no. 1, pp. 37–45, 1989.
- [4] Z. S. Xu, "Intuitionistic fuzzy aggregation operators," *IEEE Transactions on Fuzzy Systems*, vol. 15, no. 6, pp. 1179–1187, 2007.
- [5] Z. Xu and R. R. Yager, "Some geometric aggregation operators based on intuitionistic fuzzy sets," *International Journal of General Systems*, vol. 35, no. 4, pp. 417–433, 2006.
- [6] Z. Xu and R. R. Yager, "Dynamic intuitionistic fuzzy multi-attribute decision making," *International Journal of Approximate Reasoning*, vol. 48, no. 1, pp. 246–262, 2008.
- [7] G. W. Wei, "Some geometric aggregation functions and their application to dynamic multiple attribute decision making in intuitionistic fuzzy setting," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 17, no. 2, pp. 179–196, 2009.
- [8] G. Wei, "Some induced geometric aggregation operators with intuitionistic fuzzy information and their application to group decision making," *Applied Soft Computing*, vol. 10, no. 2, pp. 423–431, 2010.
- [9] G. Wei and X. Zhao, "Some induced correlated aggregating operators with intuitionistic fuzzy information and their application to multiple attribute group decision making," *Expert Systems with Applications*, vol. 39, no. 2, pp. 2026–2034, 2012.
- [10] X. Zhao and G. Wei, "Some intuitionistic fuzzy Einstein hybrid aggregation operators and their application to multiple attribute decision making," *Knowledge-Based Systems*, vol. 37, pp. 472–479, 2013.
- [11] Z. Xu and M. Xia, "Induced generalized intuitionistic fuzzy operators," *Knowledge-Based Systems*, vol. 24, no. 2, pp. 197–209, 2011.
- [12] Z. Xu and Q. Chen, "A multi-criteria decision making procedure based on interval-valued intuitionistic fuzzy bonferoni means," *Journal of Systems Science and Systems Engineering*, vol. 20, no. 2, pp. 217–228, 2011.
- [13] F. Jin, L. Pei, H. Chen, and L. Zhou, "Interval-valued intuitionistic fuzzy continuous weighted entropy and its application to multi-criteria fuzzy group decision making," *Knowledge-Based Systems*, vol. 59, pp. 132–141, 2014.
- [14] X. Qi, C. Liang, and J. Zhang, "Generalized cross-entropy based group decision making with unknown expert and attribute weights under interval-valued intuitionistic fuzzy environment," *Computers & Industrial Engineering*, vol. 79, pp. 52–64, 2015.
- [15] G.-W. Wei, H.-J. Wang, and R. Lin, "Application of correlation coefficient to interval-valued intuitionistic fuzzy multiple attribute decision-making with incomplete weight information," *Knowledge and Information Systems*, vol. 26, no. 2, pp. 337–349, 2011.
- [16] Y. Tang, L.-L. Wen, and G.-W. Wei, "Approaches to multiple attribute group decision making based on the generalized dice similarity measures with intuitionistic fuzzy information," *International Journal of Knowledge-Based and Intelligent Engineering Systems*, vol. 21, no. 2, pp. 85–95, 2017.
- [17] T.-Y. Chen, "The inclusion-based TOPSIS method with interval-valued intuitionistic fuzzy sets for multiple criteria group decision making," *Applied Soft Computing*, vol. 26, pp. 57–73, 2015.
- [18] T.-Y. Chen, "An interval-valued intuitionistic fuzzy permutation method with likelihood-based preference functions and its application to multiple criteria decision analysis," *Applied Soft Computing*, vol. 42, pp. 390–409, 2016.
- [19] J.-C. Wang and T.-Y. Chen, "Likelihood-based assignment methods for multiple criteria decision analysis based on interval-valued intuitionistic fuzzy sets," *Fuzzy Optimization and Decision Making*, vol. 14, no. 4, pp. 425–457, 2015.
- [20] G.-W. Wei, "Gray relational analysis method for intuitionistic fuzzy multiple attribute decision making," *Expert Systems with Applications*, vol. 38, no. 9, pp. 11671–11677, 2011.

## Retraction

# Retracted: A Cost Allocation Decision Model for Air Pollution Control

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] D. Qin and C. Gao, "A Cost Allocation Decision Model for Air Pollution Control," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4116527, 13 pages, 2022.

## Research Article

# A Cost Allocation Decision Model for Air Pollution Control

**Dong-Sheng Qin**  and **Chang-Yuan Gao** 

*Harbin University of Science and Technology, Harbin, China*

Correspondence should be addressed to Chang-Yuan Gao; [gaocy2002@126.com](mailto:gaocy2002@126.com)

Received 1 December 2021; Accepted 29 December 2021; Published 31 January 2022

Academic Editor: Daqing Gong

Copyright © 2022 Dong-Sheng Qin and Chang-Yuan Gao. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Air pollution control as the background of a cost allocation method is based on the Shapley value to determine the core stakeholder, so fair pollution control projects and the establishment of the atmospheric pollution of governance cost allocation model are put forward for the solution of air pollution coordinated by the government supervision and the atmospheric pollution control collaborative group. The results show that the cost allocation model of air pollution control based on Shapley value is more reasonable, and the cost of stakeholders is reduced to a certain extent, and the risk of the participants is reduced so that it maximizes social benefits.

## 1. Introduction

Over the past 30 years of the Reform and Opening-up policy, China has made some progress in air pollution prevention and control. A part of single atmospheric pollutants has reached the emission standard step by step. However, new pollution caused by the mixing of multiple pollution sources has attracted widespread attention with the complexity of production and living activities. PM<sub>2.5</sub> or haze is formed by a mixed reaction of various pollutants, such as gases and particulates, including volatile organic compounds (VOCs). Exposure to fine particulate matter (PM<sub>2.5</sub>) with a diameter of < 2.5 μm is a recognized cause of respiratory diseases in children [1]. At present, the generation of air pollutants is continuing to increase. Effective pollution prevention and control should decrease the increment and storage of pollutants. At present, the problem of air pollution control is still facing many difficulties and needs to be solved urgently.

In my country's urban development system, the urban development of the Beijing-Tianjin-Hebei region has driven the overall economic development of our country, but the environmental problems in these areas have gradually become the main concern of people. [2] The Beijing-Tianjin-Hebei

region, which is a key area of pollution prevention and control, suffers from grave ecological and environmental problems, especially serious air pollution and frequent haze. At the same time, the Beijing-Tianjin-Hebei region, which has a resource-dependent economy dominated by clusters of heavy industries, has long borne the highest PM<sub>2.5</sub> pollution levels in China, prompting serious concerns about the region's disease burden. [3] The overall air quality in the Beijing-Tianjin-Hebei region affects each other and cannot be prevented and controlled individually. It is urgent to strengthen regional ecological protection and construction. However, the integration of the Beijing-Tianjin-Hebei region is still at an early stage, and many systems need to be further established and refined. It is also difficult to significantly improve the quality of the atmospheric environment in the short term. Under this circumstance, coordinating the contradiction between the environmental demand and environmental capacity of the Beijing-Tianjin-Hebei region is prominent.

Noncooperative policies for transboundary pollution abatement are inefficient and ineffective, but cooperative policies for pollution abatement are seldom successful in regions that have different development goals and conflicting interests. [4] To reduce the harm of haze to the life and health of residents and accelerate the coordinated prevention and



control of air pollution in the Beijing-Tianjin-Hebei region require more resources and capital investment. However, the Beijing-Tianjin-Hebei cooperation has not clearly defined and divided rights, responsibilities, and benefits in air pollution prevention and control, especially the cost-sharing mechanism, so it is difficult to achieve effective integration.

In response to the above problems, this paper conducts a theoretical analysis on the cost allocation of the Beijing-Tianjin-Hebei cooperation in air pollution control. We propose to establish a cost-sharing model of atmospheric pollution in Beijing, Tianjin, and Hebei by using Shapley value. This model can better identify stakeholders in the cooperative governance, maximize the overall and individual interests, reduce the cost of air pollution prevention and control, and finally solve the coordination problem of air pollution control in Beijing, Tianjin, and Hebei.

## 2. Literature Review

Haze pollution not only negatively influences public health but also causes great economic losses [5, 6]. In the existing research on air pollution control, the main focus is on the causes of air pollution and the control measures of air pollution. Yang et al. quantitatively evaluated the vulnerability of the atmospheric environment in the Beijing-Tianjin-Hebei region through space-time comparison [7]. The relationship between atmospheric environmental vulnerability and exposure index, sensitivity index, and adaptability index is helpful to analyze atmospheric environmental vulnerability. The results show that the air environment vulnerability of 13 cities in the Beijing-Tianjin-Hebei region shows obvious spatial heterogeneity. Liang and Wang systematically summarized the Beijing-Tianjin-Hebei urban agglomeration implemented air pollution control mode, which is multilevel, cross regional, and multidirectional [8]. This includes a state-agglomeration-city linkage structure, multiprovincial, and cross-regional linkage governance, and multidirectional linkage mechanisms involving industrial access, energy structure, green transportation, cross-regional assistance, monitoring and early warning, consultation, and accountability to analyze the temporal and spatial characteristics by using the concentration data of six kinds of air pollutants. Sun et al. selected indicators from the three aspects of air pollution characteristics, natural condition characteristics, and trade characteristics, and used the entropy-top SIS method to determine the priority order of the JPCAP region [9]. By determining the key areas of four regions, key areas of air pollution control were identified, providing practical guidance and theoretical basis for regional joint control of air pollution. Piersanti et al. discussed the scenarios developed for 2030 in Italy's National Air Pollution Control Programme, using 2010 as the reference year, and also used these scenarios to provide a comprehensive approach to calculating the impact of the plan on health effects (mortality) and associated costs, providing an important framework and assessing measures to reduce air pollution with an integrated approach [10]. Ikeuchi et al. proposed that health risks caused by PM<sub>2.5</sub> would increase with the decrease of precipitation duration and incidence

[11]. Schwartz et al. believed that particulate air pollution at common concentrations is associated with daily death [12]. Gulia et al. highlighted the problem of high spatiotemporal variation of air pollution levels in urban areas and the methods that can be used to eliminate pollutants. The efficiency of the prototypes/devices developed using these processes has also been compared worldwide, and studies have shown that such controls would be very useful in reducing air pollution in hot spots with large spatiotemporal variability [13].

Given the formation reasons of haze and its harm to human health, Elzbieta believed that the biggest cause of smog and air pollution is the burning of garbage in the stoves; other causes include exhaust from large factories, burning coal in the stoves, and automobile exhaust [14]. Urbanization, industrialization, and increasing fossil fuel consumption are generally identified as the main contributors to poor air quality [15]. Zhao and Yuan innovatively used precipitation as an instrumental variable to alleviate the endogeneity of haze pollution variable and found that haze pollution seriously reduces the quality of China's economic development. The loss of labor supply, deurbanization, and interruption of human capital are the three transmission channels through which haze pollution affects the quality of China's economic development [16]. Wang et al. studied the changes in the physical and chemical properties of air particles collected in Beijing during the typical transition from haze to dust [17]. The study found that detailed information on the physical and chemical properties of atmospheric particles in the typical heavy pollution process could be provided in a short time, and such short-term changes should be taken into account to more accurately assess the environmental, climatic, and health impacts of airborne particulate matter. Intense economic and human activities in megacities lead to air pollution emissions, resulting in high concentrations of air pollutants in the atmosphere that harm human health, cause regional haze and acid deposition, damage crops, affect regional air quality, and contribute to climate change [18]. Zhang et al. studied the diurnal variation of atmospheric boundary layer height, analyzed temperature, wind direction and vertical structure of wind, and characteristics of velocity in the boundary layer; the weak vertical exchange between the boundary layer and the upper layer will promote the formation of fog and haze [19].

SO<sub>x</sub>, NO<sub>x</sub>, CO<sub>x</sub>, volatile organic compounds (VOCs), and Hg vapor directly and indirectly harm the atmospheric environment and human health by contributing to the formation of photochemical smog, acid rain, and haze and posing risks of potential toxicity (e.g., carcinogenicity) [20]. Wang and Yuan tested the impact of air pollution control on ecological total factor energy efficiency and the moderating effect of ownership structure by using panel data of 37 subindustries of China's industrial sectors from 2003 to 2014. They found that high-pollution industries had a significant short-term positive impact, while medium-pollution industries had a significant short-term negative impact. Low pollution industries have a significant long-term negative impact. Finally, some concrete policy



suggestions are put forward [21]. Kim et al. determined the causal relationship between the risk perception of particulate matter and satisfaction of outdoor activities in South Korea, conducted an online survey of 412 people, and conducted confirmatory factor analysis using the structural equation model. The results showed that the perceived risk of particulate matter was higher when people had no interest or trust in public opinion or policy, which increased people's perception of health risks and, in turn, reduced their satisfaction with outdoor activities [22]. Zhao et al. studied the Beijing-Tianjin-Hebei region as an example, used the Shapley value method to reasonably distribute the model benefits, construct the optimal emission reduction plan, and determine the optimal annual decision of each province. The results show that the model reduces the total cost of SO<sub>2</sub> pollution reduction [23]. Gao et al. evaluated the risk of particulate pollution to human health by comparing samples collected from seven different functional areas in Beijing [24].

At present, haze air pollution seriously harms global human health. Many scholars at home and abroad have made detailed explanations on the sources of air pollutants, mainly from organic compounds such as SO<sub>x</sub>, NO<sub>x</sub>, and CO<sub>x</sub>, open incineration of garbage, and automobile exhaust emissions. For the study of air pollution, many scholars focus on empirical research. However, the Shapley value method is a model of the income distribution of participants in dynamic cooperative alliances. At present, the Shapely value method is mainly used to solve the income distribution in cooperative projects at home and abroad. The Shapley value method is seldom used to study the cost of air pollution control.

In this paper, the cost of air pollution control is allocated according to the Shapley value method. Compared with the existing research work, this model avoids equal distribution and other relatively simple distribution to a certain extent, which is scientific, reasonable, and fair. This paper establishes an air pollution control cost-sharing model based on the Shapley value and determines the core stakeholders to make the cost allocation more reasonable, scientific, and fair. The Shapley model reduces the charges of each stakeholder, disperses the risks of participants to a certain extent, and effectively ensures the smooth progress of the project. We found that the cost distribution model based on Shapley value can optimize the cost and maximize the social benefit. Because air pollution control needs a long process, the reduction of cost and risk can effectively ensure the completion of air pollution control, which plays an important role in promoting air pollution control.

### 3. Basic Theory of Air Pollution Control Cost Allocation

**3.1. Shapley Value Method.** Let  $I = \{1, 2, \dots, n\}$  be the total number of project participants,  $S$  be any subset of  $I$ , and  $V$  be the characteristic function defined above  $I$ , that is, the benefit of a certain cooperation  $S$ .

$$\begin{aligned} V(\emptyset) &= 0(1), \\ V(S_1 \cup S_2) &\geq V(S_1) + V(S_2), S_1 \cap S_2 = \emptyset. \end{aligned} \quad (1)$$

With  $Y_i$  said  $I$  in the case of a project  $I$  partners from the alliance profit maximum  $V(I)$  in income, namely, the income gained by the each participant in the  $I$  uses letters to represent  $Y$ , this model is a cost allocation which should not only meet the rationality of overall but also at the same time to meet the individual rationality [25, 26], as shown below:

$$\begin{aligned} \text{the overall rational: } & \sum_{i=1}^n Y_i = V(I), i = 1, 2, \dots, n, \\ \text{individual rationality: } & Y_i \geq V(i), i = 1, 2, \dots, n, \end{aligned} \quad (2)$$

where  $V(I)$  represents the benefit that  $I$ , a participant, can obtain by completing the project alone, and  $V(I)$  represents the maximum benefit that can be obtained in many cooperative alliances [27, 28]. At this point, we get the Shapley value we want to find, which is called  $Y_i$  ( $v$ ):

$$\begin{aligned} Y_i(v) &= \sum_{S_{res}} W(|S|) [V(S) - V(S \setminus i)], i = 1, 2 \dots n, \\ W(|S|) &= \frac{(n - |S|)! (|S| - 1)!}{n!}, \end{aligned} \quad (3)$$

where  $S$  is a subset of all combinations of  $I$ , and  $|S|$  is the number of project-related stakeholders in a collaboration  $S$ , and  $W(|S|)$  is the weight.  $V(S)$  refers to the cooperation benefits obtained or costs borne by this subset  $S$ , and  $V(S \setminus i)$  refers to the benefits obtained or costs borne by other members of  $S$  after the elimination of participant  $I$  [29].

**3.2. Stakeholder Theory.** Rowley, a foreign scholar, was the first to establish the stakeholder theory. In a narrow sense, a stakeholder means that an individual and a group participate in the same project, provide human, financial, and other resources in the project, bear a part of the project risks, and finally share interests [30, 31]. In the late 1990s, there emerged a method represented by Mitchell and Wood, called the "Mitchell scoring method," which mainly divided legitimacy, power, and urgency into three attributes. According to these three attributes, stakeholders can be divided into three types: the first type is the defining stakeholder, which has three attributes at the same time and is the primary object to be paid attention to. The second type is the anticipatory stakeholder, with any combination of two attributes. The third type is the potential stakeholder, in which participants have only one of the three characteristics [32].

## 4. Air Pollution Control Cost Allocation Model

**4.1. Core Stakeholders of Air Pollution Control Cost Allocation Project.** According to the participation of each participant in the process of air pollution control project and the degree of influence of each participant on air pollution control, [33] stakeholders in the air pollution control model are divided into three categories through analysis, as shown in Table 1.

In atmospheric pollution control projects, as core stakeholders are the key members of air pollution control

and also the main body to share the cost of air pollution, they play a vital role in air pollution control, so it is particularly important to determine the cost allocation among core stakeholders [34].

**4.2. Cost Allocation Model of Air Pollution Control Based on Shapley Value.** Based on the above analysis of stakeholders, a model is established for the allocation process of air pollution treatment costs. Combined with the relevant knowledge of Shapley value, hypotheses are firstly proposed for the model. The main hypotheses are as follows:

Assume that participants  $I = \{1, 2, \dots, n\}$ , including government departments of two places and enterprises of two places, are represented by A, B, C, D, E, and F, respectively. A represents government departments of Beijing, B represents enterprises of Beijing, C represents the government of Hebei, D represents enterprises of Hebei, E represents the government of Tianjin, and F represents enterprises of Tianjin as shown in Table 2.

Based on the analysis of the above characteristic functions and the Shapley value method, the cost allocation model of regional cooperative air pollution control is established, which is mainly calculated according to the following two formulas.

$$W(|S|) = \frac{(n - |S|)(|S| - 1)!}{n!}, \quad (4)$$

$$Y_i(V) = \sum_{S_i \in S} W(|S|)[V(S) - V(S/i)], \quad i = 1, 2, \dots, n. \quad (5)$$

**4.3. Cost Allocation of Beijing Government Departments in the Air Pollution Control Project Alliance.** The values in the table are calculated according to formula (4) and formula (5). The constituent elements in the subset S represent the stakeholders involved in the project. They, respectively, say that the cooperative alliance  $V(S)$  participated by Beijing government departments represents the air pollution control costs borne by the cooperative alliance participated by the Beijing municipal government  $V(S/I)$  represents the removal of Beijing city. After the government, this cooperative alliance would have borne the cost and  $[V(S) - V(S/i)]$  refers to the cost borne by the project with the participation of Beijing municipal government minus the cost borne by the project without the participation of Beijing municipal government.  $W(|S|)$  refers to the proportion of this cooperation alliance in cooperation with the participation of Beijing municipal government. The sum of  $W(|S|)$  is 1.  $W(|S|)[V(S) - V(S/i)]$  represents the Shapley value of Beijing municipal Government in this cooperative alliance, from which we can get the best cost allocation of Beijing municipal government. According to the formula, the final cost that Beijing municipal government should bear is as shown in Table 3.

$$\begin{aligned} Y_i(V) = & \frac{1}{6} V(S1) + \frac{1}{30} [V(S5) - V(S2)] + \frac{1}{30} [V(S7) - V(S4)] + \frac{1}{30} [V(S9) - V(S3)] \\ & + \frac{1}{30} [V(S7) - V(S4)] + \frac{1}{30} [V(S9) - V(S3)] + \frac{1}{60} [V(S11) - V(S6)] + \frac{1}{60} [V(S13) - V(S8)] \\ & + \frac{1}{60} [V(S14) - V(S10)] + \frac{1}{60} [V(S15) - V(S12)] + \frac{1}{30} [V(S19) - V(S17)] + \frac{1}{30} [V(S23) - V(S16)] \\ & + \frac{1}{60} [V(S31) - V(S21)] + \frac{1}{60} [V(S32) - V(S20)] + \frac{1}{60} [V(S33) - V(S18)] + \frac{1}{60} [V(S35) - V(S24)] \\ & + \frac{1}{60} [V(S37) - V(S26)] + \frac{1}{60} [V(S38) - V(S22)] + \frac{1}{60} [V(S40) - V(S25)] + \frac{1}{60} [V(S43) - V(S34)] \\ & + \frac{1}{60} [V(S45) - V(S36)] + \frac{1}{60} [V(S49) - V(S39)] + \frac{1}{60} [V(S50) - V(S27)] + \frac{1}{60} [V(S51) - V(S29)] \\ & + \frac{1}{60} [V(S52) - V(S41)] + \frac{1}{60} [V(S53) - V(S28)] + \frac{1}{60} [V(S54) - V(S42)] + \frac{1}{60} [V(S55) - V(S30)] \\ & + \frac{1}{30} [V(S58) - V(S44)] + \frac{1}{30} [V(S59) - V(S48)] + \frac{1}{30} [V(S60) - V(S46)] + \frac{1}{30} [V(S61) - V(S47)] \\ & + \frac{1}{30} [V(S62) - V(S56)] + \frac{1}{6} [V(S63) - V(S57)]. \end{aligned} \quad (6)$$

TABLE 1: Classification of stakeholders of air pollution control projects.

Stakeholders	Institutions	Mechanism classification basis
Core stakeholders	Local governments and enterprises in Beijing, Tianjin and Hebei	The indispensable participants of the air pollution control project have an inseparable interest with the control and cost sharing of the project
General stakeholders	Beijing-Tianjin-Hebei and surrounding areas' air pollution prevention and control collaborative group	Participants who are closely connected with air pollution control projects have a certain degree of project participation and bear certain risks
Marginal stakeholders	The central government	It has a little influence on air pollution control projects and does not directly contact each participant

TABLE 2: Making assumptions about the model.

A	B	C	D	E	F	Subset	Function	Number of participants
Participate in						$S_1$	$V(S_1)$	1
	Participate in					$S_2$	$V(S_2)$	1
		Participate in				$S_3$	$V(S_3)$	1
			Participate in			$S_4$	$V(S_4)$	1
Participate in	Participate in					$S_5$	$V(S_5)$	2
		Participate in	Participate in			$S_6$	$V(S_6)$	2
Participate in			Participate in			$S_7$	$V(S_7)$	2
	Participate in	Participate in				$S_8$	$V(S_8)$	2
Participate in		Participate in				$S_9$	$V(S_9)$	2
	Participate in		Participate in			$S_{10}$	$V(S_{10})$	2
Participate in		Participate in	Participate in			$S_{11}$	$V(S_{11})$	3
	Participate in	Participate in	Participate in			$S_{12}$	$V(S_{12})$	3
Participate in	Participate in	Participate in				$S_{13}$	$V(S_{13})$	3
Participate in	Participate in		Participate in			$S_{14}$	$V(S_{14})$	3
Participate in	Participate in	Participate in	Participate in			$S_{15}$	$V(S_{15})$	4
				Participate in		$S_{16}$	$V(S_{16})$	1
					Participate in	$S_{17}$	$V(S_{17})$	1
				Participate in	Participate in	$S_{18}$	$V(S_{18})$	2
Participate in					Participate in	$S_{19}$	$V(S_{19})$	2
	Participate in				Participate in	$S_{20}$	$V(S_{20})$	2
	Participate in			Participate in		$S_{21}$	$V(S_{21})$	2
			Participate in	Participate in		$S_{22}$	$V(S_{22})$	2
Participate in			Participate in	Participate in		$S_{23}$	$V(S_{23})$	2
		Participate in		Participate in		$S_{24}$	$V(S_{24})$	2
		Participate in			Participate in	$S_{25}$	$V(S_{25})$	2
		Participate in			Participate in	$S_{26}$	$V(S_{26})$	2
		Participate in	Participate in	Participate in		$S_{27}$	$V(S_{27})$	3
		Participate in	Participate in			$S_{28}$	$V(S_{28})$	3
		Participate in		Participate in	Participate in	$S_{29}$	$V(S_{29})$	3
			Participate in	Participate in	Participate in	$S_{30}$	$V(S_{30})$	3
Participate in	Participate in			Participate in		$S_{31}$	$V(S_{31})$	3
Participate in	Participate in				Participate in	$S_{32}$	$V(S_{32})$	3
Participate in					Participate in	$S_{33}$	$V(S_{33})$	3
	Participate in				Participate in	$S_{34}$	$V(S_{34})$	3
Participate in		Participate in			Participate in	$S_{35}$	$V(S_{35})$	3
	Participate in		Participate in		Participate in	$S_{36}$	$V(S_{36})$	3
Participate in		Participate in			Participate in	$S_{37}$	$V(S_{37})$	3
Participate in			Participate in	Participate in		$S_{38}$	$V(S_{38})$	3
	Participate in	Participate in		Participate in		$S_{39}$	$V(S_{39})$	3
Participate in			Participate in		Participate in	$S_{40}$	$V(S_{40})$	3
	Participate in	Participate in			Participate in	$S_{41}$	$V(S_{41})$	3
	Participate in		Participate in	Participate in		$S_{42}$	$V(S_{42})$	3
Participate in	Participate in			Participate in	Participate in	$S_{43}$	$V(S_{43})$	4
		Participate in	Participate in	Participate in	Participate in	$S_{44}$	$V(S_{44})$	4
Participate in	Participate in		Participate in		Participate in	$S_{45}$	$V(S_{45})$	4
	Participate in	Participate in		Participate in	Participate in	$S_{46}$	$V(S_{46})$	4
	Participate in	Participate in	Participate in		Participate in	$S_{47}$	$V(S_{47})$	4

TABLE 2: Continued.

A	B	C	D	E	F	Subset	Function	Number of participants
	Participate in		Participate in	Participate in	Participate in	$S_{48}$	$V(S_{48})$	4
Participate in	Participate in			Participate in		$S_{49}$	$V(S_{49})$	4
Participate in		Participate in	Participate in	Participate in		$S_{50}$	$V(S_{50})$	4
Participate in		Participate in		Participate in	Participate in	$S_{51}$	$V(S_{51})$	4
Participate in	Participate in	Participate in			Participate in	$S_{52}$	$V(S_{52})$	4
Participate in		Participate in	Participate in		Participate in	$S_{53}$	$V(S_{53})$	4
Participate in	Participate in		Participate in	Participate in		$S_{54}$	$V(S_{54})$	4
Participate in			Participate in	Participate in	Participate in	$S_{55}$	$V(S_{55})$	4
	Participate in	Participate in	Participate in	Participate in		$S_{56}$	$V(S_{56})$	4
	Participate in	Participate in	Participate in	Participate in	Participate in	$S_{57}$	$V(S_{57})$	5
Participate in		Participate in	Participate in	Participate in	Participate in	$S_{58}$	$V(S_{58})$	5
Participate in	Participate in		Participate in	Participate in	Participate in	$S_{59}$	$V(S_{59})$	5
Participate in	Participate in	Participate in		Participate in	Participate in	$S_{60}$	$V(S_{60})$	5
Participate in	Participate in	Participate in	Participate in		Participate in	$S_{61}$	$V(S_{61})$	5
Participate in	Participate in	Participate in	Participate in	Participate in		$S_{62}$	$V(S_{62})$	5
Participate in	Participate in	Participate in	Participate in	Participate in	Participate in	$S_{63}$	$V(S_{63})$	6

TABLE 3: Expenses borne by Beijing government departments in air pollution control projects.

A subset of S	$V(S)$	$V(S/i)$	$[V(S) - V(S/i)]$	$W( S )$	$W( S )[V(S) - V(S/i)]$
S1	$V(S1)$	0	$V(S1)$	1\6	$1/46V(S1)$
S5	$V(S5)$	$V(S2)$	$V(S5) - V(S2)$	1\30	$1/30[V(S5) - V(S2)]$
S7	$V(S7)$	$V(S4)$	$V(S7) - V(S4)$	1\30	$1/30[V(S7) - V(S4)]$
S9	$V(S9)$	$V(S3)$	$V(S9) - V(S3)$	1\30	$1/30[V(S9) - V(S3)]$
S11	$V(S11)$	$V(S6)$	$V(S11) - V(S6)$	1\60	$1/60[V(S11) - V(S6)]$
S13	$V(S13)$	$V(S8)$	$V(S13) - V(S8)$	1\60	$1/60[V(S13) - V(S8)]$
S14	$V(S14)$	$V(S10)$	$V(S14) - V(S10)$	1\60	$1/60[V(S14) - V(S10)]$
S15	$V(S15)$	$V(S12)$	$V(S15) - V(S12)$	1\60	$1/60[V(S15) - V(S12)]$
S19	$V(S19)$	$V(S17)$	$V(S19) - V(S17)$	1\30	$1/30[V(S19) - V(S17)]$
S23	$V(S23)$	$V(S16)$	$V(S23) - V(S16)$	1\30	$1/30[V(S23) - V(S16)]$
S31	$V(S31)$	$V(S21)$	$V(S31) - V(S21)$	1\60	$1/60[V(S31) - V(S21)]$
S32	$V(S32)$	$V(S20)$	$V(S32) - V(S20)$	1\60	$1/60[V(S32) - V(S20)]$
S33	$V(S33)$	$V(S18)$	$V(S33) - V(S18)$	1\60	$1/60[V(S33) - V(S18)]$
S35	$V(S35)$	$V(S24)$	$V(S35) - V(S24)$	1\60	$1/60[V(S35) - V(S24)]$
S37	$V(S37)$	$V(S26)$	$V(S37) - V(S26)$	1\60	$1/60[V(S37) - V(S26)]$
S38	$V(S38)$	$V(S22)$	$V(S38) - V(S22)$	1\60	$1/60[V(S38) - V(S22)]$
S40	$V(S40)$	$V(S25)$	$V(S40) - V(S25)$	1\60	$1/60[V(S40) - V(S25)]$
S43	$V(S43)$	$V(S34)$	$V(S43) - V(S34)$	1\60	$1/60[V(S43) - V(S34)]$
S45	$V(S45)$	$V(S36)$	$V(S45) - V(S36)$	1\60	$1/60[V(S45) - V(S36)]$
S49	$V(S49)$	$V(S39)$	$V(S49) - V(S39)$	1\60	$1/60[V(S49) - V(S39)]$
S50	$V(S50)$	$V(S27)$	$V(S50) - V(S27)$	1\60	$1/60[V(S50) - V(S27)]$
S51	$V(S51)$	$V(S29)$	$V(S51) - V(S29)$	1\60	$1/60[V(S51) - V(S29)]$
S52	$V(S52)$	$V(S41)$	$V(S52) - V(S41)$	1\60	$1/60[V(S52) - V(S41)]$
S53	$V(S53)$	$V(S28)$	$V(S53) - V(S28)$	1\60	$1/60[V(S53) - V(S28)]$
S54	$V(S54)$	$V(S42)$	$V(S54) - V(S42)$	1\60	$1/60[V(S54) - V(S42)]$
S55	$V(S55)$	$V(S30)$	$V(S55) - V(S30)$	1\60	$1/60[V(S55) - V(S30)]$
S58	$V(S58)$	$V(S44)$	$V(S58) - V(S44)$	1\30	$1/30[V(S58) - V(S44)]$
S59	$V(S59)$	$V(S48)$	$V(S59) - V(S48)$	1\30	$1/30[V(S59) - V(S48)]$
S60	$V(S60)$	$V(S46)$	$V(S60) - V(S46)$	1\30	$1/30[V(S60) - V(S46)]$
S61	$V(S61)$	$V(S47)$	$V(S61) - V(S47)$	1\30	$1/30[V(S61) - V(S47)]$
S62	$V(S62)$	$V(S56)$	$V(S62) - V(S56)$	1\30	$1/30[V(S62) - V(S56)]$
S63	$V(S63)$	$V(S57)$	$V(S63) - V(S57)$	1\6	$1/6[V(S63) - V(S57)]$

According to the formula, the final cost that Beijing enterprises should bear is

$$\begin{aligned}
Y_i(V) = & \frac{1}{6}V(S2) + \frac{1}{30}[V(S5) - V(S1)] + \frac{1}{30}[V(S10) - V(S4)] + \frac{1}{60}[V(S12) - V(S6)] + \frac{1}{60}[V(S13) - V(S9)] + \frac{1}{60}[V(S14) - V(S7)] \\
& + \frac{1}{60}[V(S15) - V(S11)] + \frac{1}{30}[V(S20) - V(S17)] + \frac{1}{30}[V(S20) - V(S17)] + \frac{1}{30}[V(S22) - V(S16)] + \frac{1}{60}[V(S31) - V(S5)] \\
& + \frac{1}{60}[V(S32) - V(S19)] + \frac{1}{60}[V(S34) - V(S18)] + \frac{1}{60}[V(S36) - V(S25)] + \frac{1}{60}[V(S39) - V(S24)] + \frac{1}{60}[V(S41) - V(S26)] \\
& + \frac{1}{60}[V(S42) - V(S22)] + \frac{1}{60}[V(S43) - V(S33)] + \frac{1}{60}[V(S45) - V(S40)] + \frac{1}{60}[V(S46) - V(S29)] + \frac{1}{60}[V(S47) - V(S28)] \\
& + \frac{1}{60}[V(S48) - V(S30)] + \frac{1}{60}[V(S49) - V(S35)] + \frac{1}{60}[V(S52) - V(S - 37)] + \frac{1}{60}[V(S54) - V(S38)] + \frac{1}{60}[V(S56) - V(S27)] \\
& + \frac{1}{30}[V(S57) - V(S44)] + \frac{1}{30}[V(S59) - V(S55)] + \frac{1}{30}[V(S60) - V(S51)] + \frac{1}{30}[V(S61) - V(S53)] + \frac{1}{30}[V(S62) - V(S50)] \\
& + \frac{1}{6}[V(S63) - V(S58)].
\end{aligned} \tag{7}$$

According to the formula, the final cost that Hebei government should bear is

$$\begin{aligned}
Y_i(V) = & \frac{1}{6}V(S3) + \frac{1}{30}[V(S6) - V(S4)] + \frac{1}{30}[V(S8) - V(S2)] + \frac{1}{30}[V(S9) - V(S1)] + \frac{1}{60}[V(S11) - V(S7)] + \frac{1}{60}[V(S12) - V(S8)] \\
& + \frac{1}{60}[V(S13) - V(S5)] + \frac{1}{60}[V(S15) - V(S14)] + \frac{1}{30}[V(S20) - V(S17)] + \frac{1}{30}[V(S24) - V(S16)] + \frac{1}{60}[V(S26) - V(S17)] \\
& + \frac{1}{60}[V(S27) - V(S22)] + \frac{1}{60}[V(S28) - V(S25)] + \frac{1}{60}[V(S29) - V(S18)] + \frac{1}{60}[V(S35) - V(S25)] + \frac{1}{60}[V(S37) - V(S19)] \\
& + \frac{1}{60}[V(S39) - V(S21)] + \frac{1}{60}[V(S41) - V(S20)] + \frac{1}{60}[V(S44) - V(S30)] + \frac{1}{60}[V(S46) - V(S34)] + \frac{1}{60}[V(S47) - V(S36)] \\
& + \frac{1}{60}[V(S49) - V(S31)] + \frac{1}{60}[V(S50) - V(S38)] + \frac{1}{60}[V(S51) - V(S - 33)] + \frac{1}{60}[V(S52) - V(S32)] + \frac{1}{60}[V(S53) - V(S27)] \\
& + \frac{1}{30}[V(S56) - V(S44)] + \frac{1}{30}[V(S57) - V(S40)] + \frac{1}{30}[V(S58) - V(S55)] + \frac{1}{30}[V(S60) - V(S43)] + \frac{1}{30}[V(S62) - V(S45)] \\
& + \frac{1}{6}[V(S63) - V(S59)].
\end{aligned} \tag{8}$$

According to the formula, the final cost that Hebei enterprises should bear is

$$\begin{aligned}
Y_i(V) = & \frac{1}{6}V(S4) + \frac{1}{30}[V(S6) - V(S3)] + \frac{1}{30}[V(S7) - V(S1)] + \frac{1}{30}[V(S10) - V(S2)] + \frac{1}{60}[V(S11) - V(S9)] \\
& + \frac{1}{60}[V(S12) - V(S8)] + \frac{1}{60}[V(S14) - V(S5)] + \frac{1}{60}[V(S15) - V(S13)] + \frac{1}{30}[V(S22) - V(S16)] \\
& + \frac{1}{30}[V(S26) - V(S14)] + \frac{1}{60}[V(S27) - V(S26)] + \frac{1}{60}[V(S29) - V(S18)] + \frac{1}{60}[V(S35) - V(S20)] \\
& + \frac{1}{60}[V(S37) - V(S23)] + \frac{1}{60}[V(S39) - V(S19)] + \frac{1}{60}[V(S41) - V(S21)] + \frac{1}{60}[V(S43) - V(S29)] \\
& + \frac{1}{60}[V(S44) - V(S32)] + \frac{1}{60}[V(S46) - V(S41)] + \frac{1}{60}[V(S47) - V(S34)] + \frac{1}{60}[V(S49) - V(S35)] \\
& + \frac{1}{60}[V(S52) - V(S37)] + \frac{1}{60}[V(S53) - V(S31)] + \frac{1}{60}[V(S54) - V(S33)] + \frac{1}{60}[V(S55) - V(S39)] \\
& + \frac{1}{60}[V(S56) - V(S46)] + \frac{1}{30}[V(S57) - V(S51)] + \frac{1}{30}[V(S58) - V(S43)] + \frac{1}{30}[V(S59) - V(S52)] \\
& + \frac{1}{30}[V(S61) - V(S49)] + \frac{1}{30}[V(S62) - V(S60)] + \frac{1}{6}[V(S25) - V(S17)].
\end{aligned} \tag{9}$$

According to the formula, the final cost that Tianjin government should bear is

$$\begin{aligned}
Y_i(V) = & \frac{1}{6}V(S16) + \frac{1}{30}[V(S18) - V(S17)] + \frac{1}{30}[V(S21) - V(S2)] + \frac{1}{30}[V(S22) - V(S3)] + \frac{1}{30}[V(S23) - V(S6)] \\
& + \frac{1}{60}[V(S24) - V(S26)] + \frac{1}{60}[V(S27) - V(S25)] + \frac{1}{60}[V(S29) - V(S5)] + \frac{1}{30}[V(S30) - V(S19)] \\
& + \frac{1}{30}[V(S31) - V(S20)] + \frac{1}{60}[V(S27) - V(S26)] + \frac{1}{60}[V(S39) - V(S41)] + \frac{1}{60}[V(S42) - V(S28)] \\
& + \frac{1}{60}[V(S35) - V(S32)] + \frac{1}{60}[V(S38) - V(S28)] + \frac{1}{60}[V(S41) - V(S21)] + \frac{1}{60}[V(S43) - V(S36)] \\
& + \frac{1}{60}[V(S44) - V(S13)] + \frac{1}{60}[V(S46) - V(S11)] + \frac{1}{60}[V(S47) - V(S37)] + \frac{1}{60}[V(S49) - V(S14)] \\
& + \frac{1}{60}[V(S50) - V(S40)] + \frac{1}{60}[V(S51) - V(S12)] + \frac{1}{60}[V(S54) - V(S47)] + \frac{1}{60}[V(S55) - V(S33)] \\
& + \frac{1}{60}[V(S56) - V(S45)] + \frac{1}{30}[V(S57) - V(S52)] + \frac{1}{30}[V(S58) - V(S15)] + \frac{1}{30}[V(S59) - V(S61)] \\
& + \frac{1}{30}[V(S60) - V(S10)] + \frac{1}{30}[V(S62) - V(S60)] + \frac{1}{6}[V(S63) - V(S17)].
\end{aligned} \tag{10}$$

According to the formula, the final cost that Tianjin enterprises should bear is



$$\begin{aligned}
Y_i(V) = & \frac{1}{6}V(S16) + \frac{1}{30}[V(S18) - V(S16)] + \frac{1}{30}[V(S19) - V(S1)] + \frac{1}{30}[V(S20) - V(S2)] + \frac{1}{60}[V(S23) - V(S6)] \\
& + \frac{1}{60}[V(S26) - V(S24)] + \frac{1}{60}[V(S27) - V(S22)] + \frac{1}{60}[V(S28) - V(S5)] + \frac{1}{30}[V(S30) - V(S23)] \\
& + \frac{1}{30}[V(S33) - V(S10)] + \frac{1}{60}[V(S34) - V(S9)] + \frac{1}{60}[V(S36) - V(S7)] + \frac{1}{60}[V(S37) - V(S8)] \\
& + \frac{1}{60}[V(S41) - V(S27)] + \frac{1}{60}[V(S43) - V(S14)] + \frac{1}{60}[V(S44) - V(S39)] + \frac{1}{60}[V(S46) - V(S12)] \\
& + \frac{1}{60}[V(S46) - V(S42)] + \frac{1}{60}[V(S47) - V(S35)] + \frac{1}{60}[V(48) - V(S13)] + \frac{1}{60}[V(S51) - V(S11)] \\
& + \frac{1}{60}[V(S52) - V(S38)] + \frac{1}{60}[V(S53) - V(S56)] + \frac{1}{60}[V(S55) - V(S50)] + \frac{1}{30}[V(S57) - V(S42)] \\
& + \frac{1}{30}[V(S58) - V(S49)] + \frac{1}{30}[V(S59) - V(S15)] + \frac{1}{30}[V(S58) - V(S15)] + \frac{1}{30}[V(S60) - V(S62)] \\
& + \frac{1}{30}[V(S62) - V(S60)] + \frac{1}{6}[V(S63) - V(S17)].
\end{aligned} \tag{11}$$

Therefore, through the above results, we can know that the total cost of governance is the cost borne by the three governments plus the cost borne by the enterprise is  $V(S15)$ . In the air pollution control project, it is necessary to determine the total cost  $V(S15)$ , the separate treatment cost  $V(S1)$  of the government in Beijing, and the cost  $V(S2)$  of the government alone to control air pollution. Therefore, through the above results, we can know that the total cost of governance is the cost borne by the three governments plus the cost borne by the enterprise is  $V(S15)$ . In the air pollution control project, it is necessary to determine the total cost  $V(S15)$ , the separate treatment cost  $V(S1)$  of the government in Beijing, and the cost  $V(S2)$  of the government alone to control air pollution. Hebei government alone treatment cost  $V(S3)$  and Hebei enterprise independent control cost  $V(S4)$ , Tianjin government independent treatment cost  $V(S16)$ , Tianjin enterprise independent treatment cost  $V(S17)$ , by knowing the cost of the combination of the two of them, the cost of the combination of any three, the cost of the combination of any four, the combination of any five, and the cost of their joint governance, we can determine the respective costs borne by the government and enterprises in the treatment of air pollution in Beijing, Tianjin, and Hebei.

## 5. Example Analysis and Correction Based on Shapley Value

**5.1. Numerical Calculation of Distribution Model.** After collecting relevant data and sorting out the costs of air pollution control, we get the relevant costs for the treatment of air pollution in Beijing and Hebei. Taking Beijing and Hebei as examples, simplify the cost allocation model. Use A for Beijing Municipal Government, B for Beijing enterprises, C for Hebei provincial government, and D for Hebei enterprises, specific data as shown in Table 4.

According to the model established in the above section, the costs that Beijing municipal government, Beijing municipal enterprises, and Hebei provincial government should bear when

establishing a cooperative alliance to control air pollution can be calculated. The specific results are shown in Table 5.

The sum of the data in the last row of the table can get the cost that Beijing government departments should bear in the joint governance of Beijing and Hebei. The final result is 229.167 billion Yuan, which is the optimal the cost. The specific results are shown in Table 6.

The sum of the data in the last row of the table can get the cost that Beijing enterprises should bear in the joint governance of Beijing and Hebei. The final result is 82.5 billion Yuan, which is the optimal cost. The specific results are shown in Table 7.

The sum of the data in the last row of the table can get the cost that the Hebei government should bear in the joint governance of Beijing and Hebei. The final result is 224.167 billion Yuan, which is the optimal cost. The specific results are shown in Table 8.

The sum of the data in the last row of the table can get the cost that enterprises in Hebei should bear in the joint governance of Beijing and Hebei region. The final result is 86.667 billion Yuan, which is the optimal cost.

**5.2. Result Analysis.** Based on the Shapley value method, when Beijing and Hebei implement joint governance, the total cost is 622.5 billion Yuan, which can effectively solve the cooperation needs of 620 billion Yuan. Meanwhile, Beijing municipal government departments bear 229.167 billion Yuan in the cooperative governance, Beijing enterprises bear 82.5 billion yuan of governance cost, the Hebei government bears 2,241 Yuan and 6.7 billion Yuan, and the cost borne by Hebei enterprises bear 86.667 billion yuan.

Because what is allocated as the cost, the Shapley value method is generally used for the distribution of income, so its assumption condition is generally positive; that is, the income obtained is positive; then on the contrary, the cost paid is negative; that is, A bears 229.167 billion yuan in cooperative governance, B bears 82.5 billion yuan, C bears 224.167 billion yuan, and D bears 86.667 billion yuan.

TABLE 4: Calculation of air pollution control costs (unit: 100 million Yuan).

Participants	Overhead expenses
A	7600
B	3200
C	7400
D	3000
A, B	7300
C, D	7200
A, D	7100
B, C	7050
A, C	7000
B, D	7150
A, C, D	6900
B, C, D	6800
A, B, C	6500
A, B, D	6700
A, B, C, D	6200

TABLE 5: Expenses borne by Beijing Government departments in air pollution control projects (unit: 100 million Yuan).

A subset of S	V(S)	V(S/i)	$[V(S) - V(S/i)]$	W( S )	W( S )[V(S) - V(S/i)]
S1	7600	0	7600	1/4	1900.00
S5	7300	3200	4100	1/12	341.67
S7	7100	3000	4100	1/12	341.67
S9	7000	7400	-400	1/12	-33.33
S11	6900	7200	-300	1/12	-25.00
S13	6500	7050	-550	1/12	-45.83
S14	6700	7150	-450	1/12	-37.50
S15	6200	6800	-600	1/4	-150.00

TABLE 6: Expenses borne by Beijing Enterprises in air pollution control projects. (unit: 100 million Yuan).

A subset of S	V(S)	V(S/i)	$[V(S) - V(S/i)]$	W( S )	W( S )[V(S) - V(S/i)]
S2	3200	0	3200	1/4	800.00
S6	7300	7800	-500	1/12	-41.67
S7	7100	7400	-300	1/12	-25.00
S10	7150	3000	4150	1/12	345.83
S11	6900	7200	-300	1/12	-25.00
S12	6800	7000	-200	1/12	-16.67
S14	6700	7150	-450	1/12	-37.50
S15	6200	6900	-700	1/4	-175.00

TABLE 7: Expenses borne by the Hebei Government in air pollution control projects. (unit: 100 million Yuan).

A subset of S	V(S)	V(S/i)	$[V(S) - V(S/i)]$	W( S )	W( S )[V(S) - V(S/i)]
S3	7400	0	7400	1/4	1850.00
S6	7200	3000	4200	1/12	350.00
S8	7050	3200	3850	1/12	320.83
S9	7000	7600	-600	1/12	-50.00
S11	6900	7100	-200	1/12	-16.67
S12	6800	7050	-250	1/12	-20.83
S13	6500	7300	-800	1/12	-66.67
S15	6200	6700	-500	1/4	-125.00

According to the Shapley value allocation, the cooperation cost is shown in Table 9.

It can be seen from the above table that the cost borne by the cooperation is less than the cost generated by the

individual governance, which also satisfies the individual rationality. At the same time, compared with the individual governance in each region, the cost borne by each stakeholder in the joint governance is significantly less than that

TABLE 8: Expenses borne by Hebei Enterprises in air pollution control projects (unit: 100 million Yuan).

A subset of S	$V(S)$	$V(S/i)$	$[V(S) - V(S/i)]$	$W( S )$	$W( S )[V(S) - V(S/i)]$
S4	3000	0	3000	1/4	750.00
S6	7200	7400	-200	1/12	-16.67
S7	7100	7600	-500	1/12	-41.67
S10	7150	3200	3950	1/12	329.17
S11	6900	7000	-100	1/12	-8.33
S12	6800	7050	-250	1/12	-20.83
S14	6700	7300	-600	1/12	-50.00
S15	6200	6500	-300	1/4	-75.00

TABLE 9: Cost of air pollution control (unit: 100 million Yuan).

Participants	Original governance cost	Cost based on Shapley value
A	-7600	-2291.67
B	-3200	-825.00
C	-7400	-2241.67
D	-3000	-866.67
A, B	-7300	-3116.67
C, D	-7200	-3108.33
A, D	-7100	-3158.33
B, C	-7050	-3066.67
A, C	-7000	-4533.33
B, D	-7150	-1691.67
A, C, D	-6900	-5400.00
B, C, D	-6800	-3933.33
A, B, C	-6500	-5358.33
A, B, D	-6700	-5400.00

borne by the individual governance. In the absence of relevant stakeholders,  $V(\emptyset) = 0$  and  $V(S_1 \cup S_2) \geq V(S_1) + V(S_2)$  are satisfied.

To sum up, the cost obtained by using the Shapley value method is less than the cost of independent cooperation or independent governance, which can make the cost allocation more scientific, fair, and reasonable, reduce the risks borne by each stakeholder in participating in project governance, and promote the progress of the project governance. At the same time, in the actual situation, we must adjust the cost according to the actual situation and pay attention to risk control, to ensure the smooth completion of air pollution control.

5.3. *Modifying the Model.* Through the above analysis, the four benefit distribution factors, namely, cost bearing, risk bearing, contribution degree, and contract execution degree, should be considered. For convenience, the influential factors are set  $T = \{T\}$ ,  $T = 1, 2, 3,$  and  $4,$  respectively, representing these four factors. And in the cooperation mode set  $S$  of participants, the modification value of the influence factor on  $t$  cost allocation of the  $i^{\text{th}}$  participant is BIT. The specific results are shown in Table 10.

According to the above table, the following coefficient matrix  $B$  can be obtained:

$$\begin{aligned}
 B_1 &= \begin{bmatrix} b_{A1} \\ b_{A2} \\ b_{C1} \\ b_{D1} \\ b_{E1} \\ b_{F1} \end{bmatrix}, \\
 B_2 &= \begin{bmatrix} b_{A2} \\ b_{B2} \\ b_{C2} \\ b_{D2} \\ b_{E2} \\ b_{F2} \end{bmatrix}, \\
 B_3 &= \begin{bmatrix} b_{A3} \\ b_{B3} \\ b_{C3} \\ b_{D3} \\ b_{E3} \\ b_{F3} \end{bmatrix}, \\
 B_4 &= \begin{bmatrix} b_{A4} \\ b_{B4} \\ b_{C4} \\ b_{D4} \\ b_{E4} \\ b_{F4} \end{bmatrix}.
 \end{aligned} \tag{12}$$

If we arrange the matrix  $B$ , we get the matrix  $B' = \{b'_{it}\}_{6 \times 4}$

In the process of atmospheric governance, several experts are needed to score the influencing factors. At this time, there will be a fractional coefficient matrix  $A = [A1, A2, A3, A4]$ , and a new coefficient matrix  $D = B^* A = [D1, D2, D3, D4, D5, D6]$  can be obtained.

TABLE 10: Correction coefficient table.

	1. (the expenses)	2. (risk taking)	3. (contribution)	4. (Degree of contract execution)
A	bA1	bA2	bA3	bA4
B	bB1	bB2	bB3	bB4
C	bC1	bC2	bC3	bC4
D	bD1	bD2	bD3	bD4
E	bE1	bE2	bE3	bE4
F	bF1	bF2	bF3	bF4

Therefore, after adjustment, the cost to be borne by each region is

$$Y'_i(V) = \sum_{S_j \in S} W(|S|)[V(S) - V(S/i)] + D_j i \quad (13)$$

$$= 1, 2 \dots n.$$

The adjusted Shapley value is closer to the reality, and more influential factors are considered so that the cost sharing is not only more objective but also to ensure the smooth implementation of the project so that the goal of atmospheric governance can be realized faster, and the cost distribution scheme is more fair, reasonable, and scientific.

## 6. Conclusion

This paper focuses on the issue across regional pollution control cost allocation. We established an air pollution control cost-sharing model based on the Shapley value and determined the core stakeholders to make the cost allocation more reasonable, scientific, and fair. The Shapley model reduces the charges of each stakeholder, disperses the risks of participants to a certain extent, and effectively ensures the smooth progress of the project. We found that the cost distribution model based on Shapley value can optimize the cost and maximize the social benefit. Because air pollution control needs a long process, the reduction of cost and risk can effectively ensure the completion of air pollution control, which plays an important role in promoting air pollution control.

The subsequent studies could consider factors such as risk, contribution rate, and the bearing capacity of each stakeholder. Determining the influencing factors according to the different situations of stakeholders can make the cost more reasonable and realistic. Besides, the governance in different regions is not the same. The prevention and control method based on the local conditions could make the cost allocation model more widely used, better solve the air pollution problem, and maximize social benefits.

## Data Availability

The data of this project were obtained through the investigation of "National Joint Research Center for Air Pollution Prevention and Control Project Topic: Regional Dynamic and High Spatial-Temporal Resolution Air Pollution Source Emission Inventory."

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References

- [1] R. Aguilera, T. Corringham, A. Gershunov, S. Leibel, and T. Benmarhnia, "Fine particles in wildfire smoke and pediatric respiratory health in California," *Pediatrics*, vol. 147, no. 4, Article ID 27128, 2021.
- [2] J. Meng, M. Wang, and X. Xuekelaiti, "Characteristics of air pollution and environmental economic efficiency in Beijing-Tianjin-Hebei and surrounding areas," *Arabian Journal of Geosciences*, vol. 14, no. 12, 2021.
- [3] L. A. Xin, B. Hz, and X. C. Tao, "Consumption-based PM 2.5-related premature mortality in the Beijing-Tianjin-Hebei region," *The Science of the Total Environment*, vol. 800, no. 15, Article ID 149575, 2021.
- [4] Y. Jiao, M. Su, C. Ji, and S. Yang, "How to design fully cooperative policies to abate transboundary air pollution between two highly asymmetric regions: an abnormal incrementalism analysis," *Journal of Cleaner Production*, vol. 278, no. 3, Article ID 124, 2021.
- [5] H. A. Yan, B. Xn, and C. Jw, "Impacts of haze pollution on China's tourism industry: a system of economic loss analysis," *Journal of Environmental Management*, vol. 295, no. 1, Article ID 113051, 2021.
- [6] C. Zhao, M. Deng, and X. Cao, "Does haze pollution damage urban innovation? Empirical evidence from China," *Environmental Science and Pollution Research*, vol. 28, no. 13, pp. 16334–16349, 2021.
- [7] Z. Yang, S. Jing, and L. Yu, "Atmospheric environment vulnerability cause analysis for the beijing-tianjin-hebei metropolitan region," *International Journal of Environmental Research and Public Health*, vol. 15, no. 1, 128 pages, 2018.
- [8] L. Liang and Z. Wang, "Control models and spatiotemporal characteristics of air pollution in the rapidly developing urban agglomerations," *International Journal of Environmental Research and Public Health*, vol. 18, no. 11, 6177 pages, 2021.
- [9] L. Sun, J. Du, and Y. Li, "A new method for dividing the scopes and priorities of air pollution control based on environmental justice," *Environmental Science and Pollution Research*, vol. 28, no. 10, pp. 12858–12869, 2021.
- [10] A. Piersanti, I. D'Elia, M. Gualtieri et al., "The Italian national air pollution control Programme: air quality, health impact and cost assessment," *Atmosphere*, vol. 12, no. 2, 196 pages, 2021.
- [11] H. Ikeuchi, M. Murakami, and S. Watanabe, "Scavenging of PM2.5 by precipitation and the effects of precipitation pattern changes on health risks related to PM2.5 in Tokyo, Japan," *Water Science and Technology*, vol. 72, no. 8, pp. 1319–1326, 2015.
- [12] J. Schwartz, F. Laden, and A. Zanobetti, "The concentration-response relation between PM(2.5) and daily deaths," *Environmental Health Perspectives*, vol. 110, no. 10, pp. 1025–1029, 2002.
- [13] S. Gulia, R. Tiwari, S. Mendiratta, S. Kaur, S. K. Goyal, and R. Kumar, "Review of scientific technology-based solutions

## Retraction

# Retracted: A Double-Layered Belief Rule Base Model for Driving Anger Detection Using Human, Vehicle, and Environment Characteristics: A Naturalistic Experimental Study

### Computational Intelligence and Neuroscience

Received 3 October 2023; Accepted 3 October 2023; Published 4 October 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] P. Wan, X. Deng, L. Yan, X. Jing, L. Peng, and X. Wang, "A Double-Layered Belief Rule Base Model for Driving Anger Detection Using Human, Vehicle, and Environment Characteristics: A Naturalistic Experimental Study," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 5698393, 17 pages, 2022.

## Research Article

# A Double-Layered Belief Rule Base Model for Driving Anger Detection Using Human, Vehicle, and Environment Characteristics: A Naturalistic Experimental Study

Ping Wan <sup>1</sup>, Xinyan Deng <sup>1</sup>, Lixin Yan <sup>1</sup>, Xiaowei Jing <sup>1</sup>, Liqun Peng <sup>1,2</sup>,  
and Xu Wang <sup>3</sup>

<sup>1</sup>School of Transportation and Logistics, East China Jiaotong University, Nanchang 330013, Jiangxi, China

<sup>2</sup>Tsinghua University Suzhou Automotive Research Institute, Suzhou 215134, Jiangsu, China

<sup>3</sup>School of Qilu Transportation, Shandong University, Jinan 250061, Shandong, China

Correspondence should be addressed to Ping Wan; pingw04@163.com

Received 1 December 2021; Accepted 4 January 2022; Published 28 January 2022

Academic Editor: Daqing Gong

Copyright © 2022 Ping Wan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

“Road rage,” namely, driving anger, has been becoming increasingly common in auto era. As “road rage” has serious negative impact on road safety, it has attracted great concern to relevant scholar, practitioner, and governor. This study aims to propose a model to effectively and efficiently detect driving anger states with different intensities for taking targeted intervening measures in intelligent connected vehicles. Forty-two private car drivers were enrolled to conduct naturalistic experiments on a predetermined and busy route in Wuhan, China, where drivers’ anger can be induced by various incentive events like weaving/cutting in line, jaywalking, and traffic congestion. Then, a data-driven model based on double-layered belief rule base is proposed according to the accumulation of the naturalistic experiments data. The proposed model can be used to effectively detect different driving anger states as a function of driver characteristics, vehicle motion, and driving environments. The study results indicate that average accuracy of the proposed model is 82.52% for all four-intensity driving anger states (none, low, medium, and high), which is 1.15%, 1.52%, 3.53%, 5.75%, and 7.42%, higher than C4.5, BPNN, NBC, SVM, and kNN, respectively. Moreover, the runtime ratio of the proposed model is superior to that of those models except for C4.5. Hence, the proposed model can be implemented in connected intelligent vehicle for detecting driving anger states in real time.

## 1. Introduction

“Road rage” is a particular driving emotion resulting from pressure or frustration from daily life or adverse traffic environments or discourteous behaviors from surrounding traffic participants [1]. The driving emotion has been becoming an increasingly common issue affecting road safety in auto era all over the world. A report from National Highway Safety Administration of US indicated that the ratio of traffic accidents because of emotional driving like road rage accounted for 9.2%~14.8% of the total [2]. In

China, another report showed that road rage brought about 17.33 million illegal acts, leading to 83,100 traffic accidents in 2015, 1.22% higher than that of 2014 [3]. As anger has a negative impact on a driver’s perception, identification, decision, and volition process, the driver will inevitably have a degraded driving performance finally [4]. Then, an angry driver is prone to make more mistakes or lapses or violations, leading to more traffic accident involvements [5]. Therefore, a driving anger detection/warning method should be designed for effective intervening to deal with road rage in connected intelligent vehicles nowadays.



*1.1. Literature Review.* To effectively and efficiently identify different emotional states based on multisource heterogeneous data, it is important to select a suitable classification method. Additionally, appropriate emotion elicitation method is also necessary as it is a precondition for emotion detection research. Hence, a brief review of related studies is shown as follows.

*1.2. Emotion Elicitation.* When a target emotion is evoked by a specific stimulating scenario, the evoking process is called emotion elicitation. At present, most of emotion elicitation has been executed in quiet laboratory. Kessous et al. [6] adopted a voice interaction with an agent through speaking a sentence in a special contextualization to elicit several common emotions like anger, fear, happiness, and sadness. Juslin and Sloboda [7] proposed a music-based anger elicitation approach by means of different combinations of valences (positive/negative) and arousals (high/low), showing that the greatest anger occurred under the context of negative music with high arousal. Besides those ordinary emotions in daily life, driving emotion elicitation has been becoming a hot research topic. Lei et al. [8] firstly selected video clips extracted from Chinese well-known films including “The Rape of Nanking” and “Fist of Fury” as the anger induction materials in driving simulator experiment to study vehicles’ velocity characteristics in anger state. Roidl et al. [9] also induced the subjects’ anger with short film clips from word-wide famous film “Schindler’s List” to study their driving performances including lateral and longitudinal velocity and acceleration in anger state. Abdu et al. [10] made the subjects recall incentive events encountered in their daily life in detail before conducting driving simulator experiment, to elicit their anger for studying their skilled driving behavior and risk taking behavior in anger state. Based on multiple networked driving simulators, Cai and Lin [11] elicited the subjects’ anger state according to mutual interactions like blocking or changing lane abruptly in front of the subject car in networked driving simulators. Danaf et al. [12] used a series of adverse events in controlled intersection to induce the subjects’ anger state in driving simulator experiments for modeling anger and aggressive driving behavior. Nevertheless, most of emotion elicitation methods in those studies are on the basis of video, music, verbal, or behavior interaction or incentive situations in virtual environment, which inevitably limit the universality of the elicitation approaches due to individual cultural background or personal preference. Moreover, the emotion elicited in inner lab is not valid as that elicited in real traffic environment, owing to some demand characteristics or social desirability [13].

*1.3. Emotion Recognition.* At present, most of emotion recognition is realized through machine learning algorithms, which can effectively deal with multisource data. Flidlund and Izard [14], at the earliest, identified several emotions including delightfulness, sadness, anger, and fear by linear discriminants based on the subjects’ facial electromyography characteristics. Wang and Gong [15] identified various driving emotions based on a factorization model using physiological features including blood volume pulse, skin conductance (SC), respiration rate

(RR), and skin temperature. Katsis et al. [16] effectively differentiated car-racing drivers’ several emotions such as high/low stress, dysphoria, and euphoria in a driving simulator environment based on decision tree and Naïve Bayesian classifier (NBC) with features of facial electromyography, electrocardiogram, respiration, and electrodermal activity (EDA). Wan et al. [17] implemented a linear discriminant model combined with receiver operating characteristic (ROC) curve analysis to identify driving anger based on five physiological features consisting of SC, RR, heart rate (HR), and relative power spectrum of beta wave ( $\beta\%$ ) and theta wave ( $\theta\%$ ) in electroencephalogram (EEG). Except for the physiological features, a driver’s operational behaviors and driving performances as well as vehicle motions have also been used for recognizing driving emotions. Malta et al. identified driver’s frustration state based on Bayesian network (BN) with EDA and brake/acceleration pedal actuation features [18]. Lanatà et al. recognized high and low driving stress through a nearest mean classifier (NMC) based on respiration activity, heart rate variability (HRV) and electrodermal response (EDR) together with steering wheel angle corrections, velocity variance [19]. Wan et al. [20] proposed a detection model of driving anger based on SVM and multivariate time series features of driving behaviors including steering wheel angle and vehicle lateral position. Danaf et al. [12] proposed a dynamic choice-latent variable model to predict driving anger level with the features of driver personality, incentive situations, and vehicle motion states acquired in driving simulator experiments. In addition, other machine learning algorithms like backpropagation neural networks (BPNN), decision tree (C4.5), and k-nearest neighbor (kNN) are often employed to recognize kinds of driving mental states [21].

So far, most of the aforementioned studies aim to recognize some usual emotions like delightfulness, sadness, stress, and frustration under lab conditions. Few studies have been conducted to recognize driving anger in real traffic environment, especially in metropolitan area of China, where road rage is a common and serious traffic psychology issue threatening road safety. Furthermore, too many physiological indicators collected by biosensors were utilized in the aforementioned studies. However, most of the current biosensors are intrusive for a driver to some extent due to various electrodes adhering to their skin surface, which may have a negative impact on their naturalistic driving performances. Additionally, most of the current studies only tackle binary discrimination of a certain mental state (e.g., fatigue or not), without more subdivision according to its intensity, which is not helpful to take target intervening measures for a specific mental state with a certain intensity. Moreover, the recognition of driving mental states in those studies is mostly based on neural network (NN), support vector machine (SVM), fuzzy logic, and Bayes network (BN), which lack effective weight allocation in input indexes and training rules, and its accuracy depends on the completeness of data samples, leading to inefficient data use. Further, those models are not suitable to be applied for recognition system requiring high instantaneity in practice when increasing input variables. For example, the conditional probability tables of BN will have exponential growth when adding a new input, which demands much more computational consumption. Last but not least, when constructing those recognition models, either

vehicle or driver data are utilized without considering the effect of driving environment on mental state whose induction is extremely dependent on a certain environment.

Aiming at that, a high-arousal driving anger elicitation approach is firstly proposed through stimulation of specific incentive situations according to naturalistic driving experiments conducted in real traffic environment, during which nonintrusive measurements like vehicle motions and driving environments, as well as driver personal attributes, will be collected. Secondly, in order to fully utilize the multisource heterogeneous data with uncertainties to effectively and efficiently recognize driving anger states, a double-layered belief rule base (BRB) approach is utilized to propose a recognition model of driving anger states with different intensity. Then, the influencing factors of driving anger could be divided into several groups according to the attribute or category of those factors. In each group, the inference process can be separately executed in an iterative way, which reduces dimensions of the BRB greatly, especially for only one layer, meaning less complicated computation and higher instantaneity. Finally, the effectiveness and efficiency of the proposed model are evaluated and compared with other widely used classification algorithms.

## 2. Experimental Designs

**2.1. Scene Design.** In order to make every subject generate anger as much as possible, a special test route containing many busy sections across Wuchang and Hankou District of Wuhan was designed for naturalistic driving experiments conducted in this study (see Figure 1(a)). The test route was almost 50.8 km, which included two tunnels, two expressways, three central business districts, 42 signalized intersections, and 59 zebra crossings. When conducting the experiment, every subject would often confront incentive environmental events or situations, such as weaving/cutting in line, jaywalking/motorcycle occupying road, traffic jams, or traffic light waiting, which randomly and frequently happened, especially during morning and evening rush hours (see Figures 1(b)–1(d)). Hence, each subject was required to begin the experiment at about 8:00 a.m. or 5:00 p.m. In order to induce more anger further, every subject was promised to get extra paid if they could complete the experiment ahead of reference time. Note that no speeding was permitted during the whole experiment. Simultaneously, to induce anger as real as possible, the driving experiments are designed to be single-blind, namely, naturalistic, which means that every subject would conduct the whole experiment alone without any interference, according to their own driving style in daily life.

**2.2. Participants.** Forty-two private car drivers with valid C1 license were enrolled from Wuhan, a fast-growing central metropolis in China, to conduct naturalistic experiments. In order to maximize the statistical power, much more male subjects were chosen in this study, as males are more prone to be involved in angry driving than females [22]. The average age of the subjects was 37.8 (SD = 5.4) years, while their

average driving age was 9.8 (SD = 4.6) years. The population distribution of the subjects with different demographic characteristics is shown in Table 1. It is noted that every subject should be checked to be in good mental condition through a profile of mood states (POMS) questionnaire, which is important to verify the effectiveness of the driving anger elicitation approach proposed in this study. Additionally, an observer (expert) with rich driving experience (>20 years) in driving behavior field was recruited to assess every subject's anger level according to the video replay of the whole experiment process.

**2.3. Apparatus.** Related research indicated that dangerous driving behaviors such as fatigue, anger, and distraction can be identified according to lateral and longitudinal vehicle motions including yaw rate, acceleration (longitudinal), and lane departure [23, 24]. Hence, a car (see Figure 2(a)) instrumented with Mobileye C2-270 system (see Figure 2(b)) and Inertial Navigation System RT2500 (see Figure 2(c)) was used as the experimental vehicle for conducting naturalistic driving experiments. It is noted that Mobileye C2-270 system is a collision warning system that can collect time headway and lane departure data, while Inertial Navigation System RT2500 is a precise vehicle motion acquisition system that can collect the heading, pitch, and roll motion characteristics of vehicle movements like yaw angular acceleration rate and longitudinal acceleration. Moreover, a clock was equipped on dash board of the test car to remind every subject of the remaining time of the experiment.

Additionally, three high-definition cameras with resolution ratio of 800 \* 640 and frame rate of 30 (see Figure 2(d)) were pasted on the front windshield of the test vehicle to record every subject's physical or behavior reaction and driving environments around. In more detail, the first camera was used to record driving environments around like unfavorable traffic environments (i.e., jam and congestion) and incentive events from surrounding traffic participants (i.e., jaywalk and cutting in line). The second one was implemented to record the subject's facial (i.e., frown and straight face) or verbal (i.e., curse and name-calling) expression and head movement. The third one was utilized to roughly record the subject's fiercely operational (high frequency or large amplitude) behaviors of steering wheel, gear lever, and other physical behaviors like slapping steering wheel or frequent honking. The video replay from the three cameras will be implemented as intuitive evidence to calibrate driving anger level for every subject and the expert.

### 2.4. Experiment Procedure

- (1) *Experimental Protocol Signing*. Each subject was demanded to sign an experimental protocol with us if they agreed on the requirements and payment stated in the protocol. ① Each subject had to begin the formal experiment from 8:00 a.m. or 5:00 p.m. ② Each subject was prohibited from violating any traffic regulation like running red lights and speeding



FIGURE 1: The test route and traffic environment of field experimental system. (a) The test route (shown as red line). (b) Jaywalking collectively. (c) Electrical motorbike occupying the road. (d) Cutting in line in traffic congestion.

TABLE 1: The population distribution of the subjects with different demographic characteristics.

Gender		Age				Temperament			
Male	Female	22–29	30–39	40–49	50–59	Melancholic	Phlegmatic	Sanguineous	Choleric
33	9	9	12	12	9	6	15	12	9

(speed limit is 70 km/h in the whole city). ③ Each subject could get the basic paid of 300RMB (Chinese currency) for accomplishing the whole experiment, and they could get extra paid with 15 RMB/min if they completed the whole experiment ahead of reference time, i.e., 120 minutes, within which it was proved to be a little pressure for the test according to the results of several pretests. Especially, all subjects can get the basic paid of 300RMB without any deduction if they could not accomplish the experiment within the reference time, based on the spirit of institutional review board (IRB) and Chinese law on scientific research. ④ They had to report their anger level objectively and truthfully without hiding their true feelings in the following step of self-report after finishing the experiment as they were informed that there was no right or wrong about their self-reports.

(2) *Personal Characteristics Collection*. Every subject's personal characteristics such as age, gender, and driving years were simultaneously collected when

signing the experimental protocol, and temperament, strongly related with driving style, was acquired by Chen Huichang Temperament Inventory [25].

(3) *Adaptive Driving Practice*. Every subject conducted 10-minute driving practice to accommodate to handling performance of the experimental vehicle, in order to eliminate discomfort or tension when firstly driving the experimental vehicle.

(4) *Formal Naturalistic Experiment*. Every subject had to accomplish the experiment alone without any interference. Note that they could finish the experiment in their habitual driving style. Meanwhile, all driving environments including the unfavorable traffic environments and incentive events (behaviors) from surrounding traffic participants, the subject's facial/verbal expression, and physical behaviors were automatically recorded by the three HD cameras during the whole experiment.



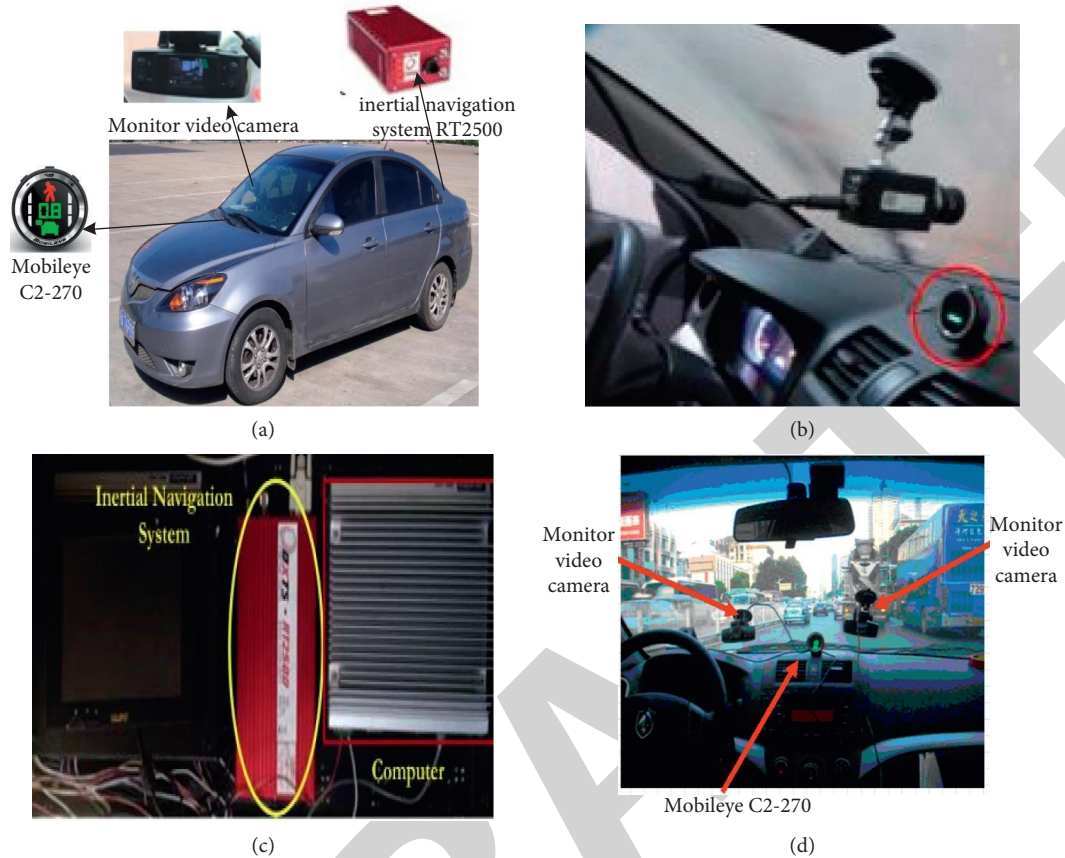


FIGURE 2: The apparatus of naturalistic experimental system. (a) Overall sketch of the apparatus system. (b) Mobileye C2-270 system. (c) Inertial Navigation RT 2500 System. (d) Monitoring video camera.

(5) *Self-Report after Experiment.* Once the experiment was completed, every subject was immediately required to recall and report their emotional type and level with a ten-point scale from 0 (not at all) to 9 (very much) every two minutes during the experimental process, or any occasion when an incentive event happened during the process, through the video replay from the three cameras in a quiet lab. Meanwhile, the observer also utilized the ten-point scale to evaluate the subject's emotional level, so as to correct some subjects' terribly subjective self-report, which had seriously deviated from the truth.

### 3. Data Reanalysis

*3.1. Driving Anger Inducing Effect.* According to the feedback from all subjects, every subject underwent several different emotions throughout the whole experimental process. Then, it is necessary to verify if the target emotion (i.e., anger) has been induced out during driving using the novel anger induction approach proposed in this study. Here, the hit rate was applied as an indicator to assess driving anger inducing effect. As each subject had to self-report their emotional type and level every two minutes and any instant an incentive driving environment (situations) occurred during driving, 2414 emotional cases in all were acquired from all subjects,

1194 of which are anger instances. As indicated in Table 2, the hit rate is 8.27% if no incentive situations occurred during driving, meaning that the extra paid for accomplishing experiment ahead of reference time did give the subjects a little time pressure. Note that the hit rate under stimulus of jaywalking/motorcycle occupying road is highest with 82.03%, while the lowest with 56.94% under the stimulus of traffic lights waiting. To sum up, the average hit rate of anger under the incentive situations reaches 74.25% (i.e.,  $(1194-75)/(2414-907)$ ), significantly higher than that under no incentive situations. Therefore, the novel anger induction approach proposed in this study is effective.

*3.2. Driving Anger Intensity Calibration.* In order to effectively and efficiently reduce the negative effect of road rage on traffic safety and efficiency, it is extremely critical to take the target warning or intervening for different driving anger states with different intensity. In this study, different anger states were calibrated based on the subjects' self-reported anger levels. It is noted that when calibrating driving anger states, any subject's self-report would be immediately adopted if the assessment discrepancy between the subject and the observer is smaller than 2, or else, three more independent experts (driving years  $\geq 25$ ) in the field of traffic psychology would be invited to reassess the subject's anger level based on the replayed video. Here, four kinds of anger

TABLE 2: Hit rate of five emotions from the scenarios with and without incentive situations.

Incentive situations	<i>Fear</i>	<i>Happy</i>	<i>Anger</i>	<i>Sad</i>	<i>Neutral</i>	<i>Total</i>	<i>Hit rate (%)</i>
No incentive situations	8	40	75	5	779	907	8.27
Jaywalking/motorcycle occupying road	44	4	388	10	27	473	82.03
Weaving/cutting in line	55	5	366	20	36	482	75.88
Traffic jam/congestion	17	8	242	27	42	336	72.02
Traffic lights waiting	12	7	123	23	51	216	56.94
Total	136	64	1194	85	935	2414	

intensity were defined in this research, namely, neutral or none anger (anger level = 0), low anger (anger level = 1, 2), mediate anger (anger level = 3, 4), and high anger (anger level  $\geq 5$ ). Consequently, 1194 anger-related samples including low anger, mediate anger, and high anger, as well as 935 neutral (none anger) samples, shown in Figure 3, were determined for the following study.

As aforementioned, some typical vehicle motions characteristics like yaw rate, acceleration, and lane departure can reflect dangerous or aggressive driving behavior. In addition, some personal characteristics including gender, age, and temperament, as well as environmental (situational) like traffic congestion and illegal traffic behaviors around, are decisive influencing factors of road rage [26]. Therefore, after calibrating the four kinds of driving anger states (modes) in terms of intensity, those corresponding vehicle characteristics, human characteristics, and environmental characteristics in line with a certain driving anger mode at that moment were all screened for constructing driving anger detection model. Then, 2129 driving samples including 1194 anger-related samples and 935 neutral (none anger) samples are partly listed in Table 3. According to the feedbacks from the participants, it is worth noting that any anger state with certain intensity would remain unchanged about 6~10 seconds until the new elicitation situation with different intensity occurs. Hence, vehicle motion signals lasting 8 (i.e., average) seconds from the moment the elicitation situation occurred were chosen for study driving anger characteristics.

## 4. Methodology

**4.1. Belief Rule Base (BRB) Inference Approach.** Yang et al. proposed an innovative inference approach, called belief rule base (BRB), to solve complicated decision-making problem [27]. The BRB approach can effectively deal with quantitative and qualitative information with uncertainty, incompleteness, subjectivity, and nonlinearity. The core of the approach is a hybrid rule base, which is established with a belief structure based on evidence theory of Dempster-Shafer, decision theory, and fuzzy inference [28]. In order to improve the approach's inference accuracy, the weights of rules and attributes as well as belief degree of output are introduced on the basis of traditional rule expression methods like fuzzy inference, during the BRB inference process using evidential reasoning (ER).

**4.1.1. Belief Rule Base Structure.** With regard to BRB inference methodology, the  $k$ th rule  $R_k$  of the rule base is expressed as follows:

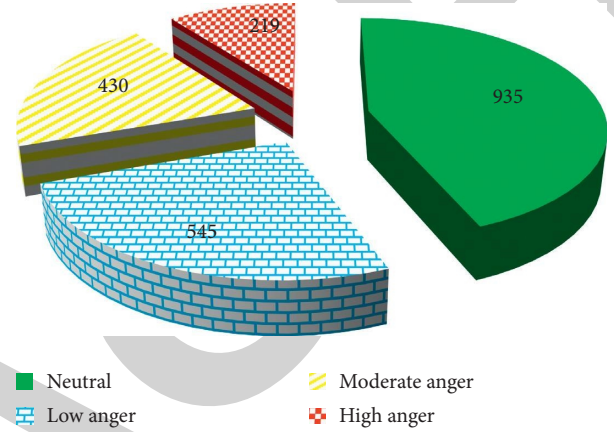


FIGURE 3: Instances of different anger states.

$$R_k: \text{ If } x_1 \text{ is } A_1^k \wedge x_2 \text{ is } A_2^k \wedge \dots \wedge x_{T_k} \text{ is } A_{T_k}^k, \text{ then } \{(D_1, \beta_{1k}), (D_2, \beta_{2k}), \dots, (D_N, \beta_{Nk})\}, \quad (1)$$

with a rule weight  $\theta_k$  and attribute weight  $\delta_{ik}$ . If the  $k$ th ( $k = 1, 2, \dots, L$ ) rule is complete, then

$$\sum_{j=1}^N \beta_{jk} = 1. \quad (2)$$

If the  $k$ th rule is incomplete, then

$$\sum_{j=1}^N \beta_{jk} < 1, \quad (3)$$

where  $x_i$  ( $i = 1, 2, \dots, T_k$ ) denotes the  $i$ th antecedent attribute (input variable) in the  $k$ th rule ( $R_k$ );  $T_k$  is the number of antecedent attributes in  $R_k$ ;  $A_i^k$  is the referential value of the  $i$ th antecedent attribute in  $R_k$ ;  $D_j$  is the  $j$ th ( $j = 1, 2, \dots, N$ ) consequent (output variable) in  $R_k$ ;  $\beta_{jk}$  is the belief degree of  $D_j$ ;  $N$  is the number of all consequents;  $\theta_k$  is the weight of  $k$ th rule;  $L$  is the number of all rules in rule base.  $\delta_{ik}$  is the weight of  $i$ th antecedent attribute in  $R_k$ .

According to equation (1), we can see that the belief if-then rule structure proposed is superior to traditional if-then rule, where the consequent is either 100% true or 100% false, leading to a strictly limited ability when representing rule (knowledge) in a real and complex world. The proposed structure can provide better flexibility in representing rule with simple or complicated, quantitative or qualitative, continuous or discrete, certain or uncertain relationship between the input and output.

TABLE 3: Partial sample data of driving anger about human, vehicle, and environment characteristics.

ID	Human			Vehicle			Environment		Driving state
	Gender	Age	Temperament	Yaw angular acceleration (deg/s <sup>2</sup> )	Longitudinal acceleration (m/s <sup>2</sup> )	Lane departure (m)	Traffic flow	Traffic behavior around	
1	Male	42	Phlegmatic	1.7001	0.6703	0.2542	Smooth	Normal	Neutral
2	Male	37	Phlegmatic	1.7611	0.8764	0.3339	Less smooth	Less severe	Neutral
3	Female	51	Phlegmatic	1.7687	1.6145	0.5007	Smooth	Severe	Low anger
4	Male	22	Choleric	2.5362	2.3043	0.7866	Less smooth	Severe	High anger
5	Female	35	Sanguineous	1.7325	0.5587	0.3176	Less smooth	Normal	Neutral
...									
2125	Female	44	Melancholic	2.1348	1.8818	0.7031	Smooth	Severe	Moderate anger
2126	Male	46	Choleric	1.8842	1.1207	0.3777	Less smooth	Less severe	Low anger
2127	Male	32	Phlegmatic	2.5575	1.4677	0.6273	Smooth	Less severe	Moderate anger
2128	Male	38	Phlegmatic	1.9329	1.5429	0.6646	Obstructive	Less severe	Low anger
2129	Male	24	Choleric	2.4033	2.3567	0.6987	Obstructive	Less severe	Moderate anger

4.1.2. *BRB Inference Using ER.* In order to compute the activation weight of  $k$ th rule  $R_k$ , the matching degree (i.e., membership) of the input  $x_i$  to the referential values of this antecedent attribute of  $R_k$  needs to be calculated firstly. The membership can be calculated by using utility-based equivalence transformation techniques, shown as follows:

$$a_{i,q}^k = \frac{(A_{i,q+1}^k - x_i)}{(A_{i,q+1}^k - A_{i,q}^k)}, \quad A_{i,q}^k < x_i \leq A_{i,q+1}^k, \quad (4)$$

$$a_{i,q+1}^k = \frac{(x_i - A_{i,q}^k)}{(A_{i,q+1}^k - A_{i,q}^k)}, \quad A_{i,q}^k < x_i \leq A_{i,q+1}^k,$$

where  $x_i$  ( $i = 1, 2, \dots, T_k$ ) denotes the  $i$ th input variable in  $R_k$ ;  $A_{i,q}^k, A_{i,q+1}^k$  are the two adjacent referential values of  $x_i$ , respectively;  $a_{i,q}^k$  is the degree to which the input  $x_i$  belongs to the referential value  $A_{i,q}^k$ , while  $a_{i,q+1}^k$  is the degree to which the input  $x_i$  belongs to the referential value  $A_{i,q+1}^k$ .

After determining  $a_{i,q}^k$ , the activation weight of the  $k$ th rule  $w_k$  is calculated as follows:

$$\beta_n = \frac{\prod_{k=1}^L (w_k \beta_{n,k} + 1 - w_k \sum_{i=1}^N \beta_{i,k}) - \prod_{k=1}^L (1 - w_k \sum_{i=1}^N \beta_{i,k})}{\sum_{j=1}^N \prod_{k=1}^L (w_k \beta_{j,k} + 1 - w_k \sum_{i=1}^N \beta_{i,k}) - (N-1) \prod_{k=1}^L (1 - w_k \sum_{i=1}^N \beta_{i,k}) - \prod_{k=1}^L (1 - w_k)} \quad (7)$$

After  $\beta_n$  is determined, the driving anger states with different intensity can be calculated as follows:

$$\text{Anger}(Y) = D_1 \beta_1 + D_2 \beta_2 + \dots + D_N \beta_N. \quad (8)$$

As the driving anger states defined in this study are comprised of neutral, low anger, moderate anger, and high

$$w_k = \frac{\theta_k \prod_{i=1}^{T_k} (\alpha_{i,q}^k)^{\bar{\delta}_i}}{\sum_{l=1}^L \left[ \theta_l \prod_{i=1}^{T_l} (\alpha_{i,q}^l)^{\bar{\delta}_i} \right]}, \quad (5)$$

where  $\bar{\delta}_i = \delta_i / \max_{i=1, \dots, T_k} \{\delta_i\}$ . If  $\bar{\delta}_i = 0$ , then  $(\alpha_{i,q}^k)^{\bar{\delta}_i} = 1$ ; it indicates that the  $i$ th antecedent attribute does not have any impact on activation weight of  $R_k$ . If  $\bar{\delta}_i = 1$ , then  $(\alpha_{i,q}^k)^{\bar{\delta}_i} = \alpha_{i,q}^k$ ,

It indicates that the  $i$ th antecedent attribute has the biggest impact on activation weight of  $R_k$ .

Given that the input is denoted by  $X$ , the outcome of BRB inference is then denoted as follows:

$$Y(x) = \{(D_j, \beta_j), j = 1, 2, 3, \dots, N\}. \quad (6)$$

Equation (6) can be interpreted as the consequent is  $D_1$  with a belief degree of  $\beta_1$ ,  $D_2$  with a belief degree of  $\beta_2, \dots$ , and  $D_N$  with a belief degree of  $\beta_N$ . Based on the analytical format of ER approach, the combined belief degree  $\beta_j$  can be calculated as follows:

anger, in terms of intensity, Anger (Y) can also be expressed as follows:

$$\text{Anger}(Y) = D_1 \beta_1 + D_2 \beta_2 + D_3 \beta_3 + D_4 \beta_4. \quad (9)$$

Hence, Anger (Y) is a continuous value within the range of  $[0, 3]$ . In order to directly compare the estimated output of



the proposed model and the observed output from the subjects' self-reports, it is necessary to discretize Anger ( $Y$ ), shown as follows:

$$\text{Estimated\_Anger}(y) = \begin{cases} 0 & 0 \leq \text{Anger}(y) \leq 0.5, \\ 1 & 0.5 < \text{Anger}(y) \leq 1.5, \\ 2 & 1.5 < \text{Anger}(y) \leq 2.5, \\ 3 & 2.5 < \text{Anger}(y) \leq 3. \end{cases} \quad (10)$$

**4.1.3. Training of BRB Model.** The initial belief rule base including BRB model parameters comprised of  $\theta_k$ ,  $\beta_{jk}$  and  $\delta_{ik}$  for the  $k$ th rule ( $R_k$ ) can be generated randomly or specified by famous experts in the fields of driving behavior or traffic psychology. Correspondingly, the initial belief rule base is inaccurate to some extent because of the experts' subjectivity or arbitrariness. Therefore, the initial belief rule base needs to be trained for optimization. In this study, abundant historical samples data collected in the naturalistic experiments were utilized to train the initial belief rule base for better representing actual rules (knowledge). The training process is illustrated as shown in Figure 4.

As indicated in Figure 4,  $\widehat{x}_m$  ( $m = 1, 2, 3, \dots, M$ ) is the input of the real system and the BRB system proposed in this study,  $\widehat{y}_m$  is the observed output of the real system,  $y_m$  is the estimated output of the BRB system, and  $\xi(p)$  is the discrepancy between the observed output and the estimated output. The objective of the training process is to get the minimum of  $\xi(p)$ ; i.e., the objective function can be defined as follows:

$$\min \xi(p) = \frac{\sum_{m=1}^M (y_m - \widehat{y}_m)^2}{M}, \quad (11)$$

subject to

$$0 \leq \beta_{jk}, \theta_k, \delta_{ik} \leq 1. \quad (12)$$

where  $p = p(\beta_{jk}, \theta_k, \delta_{ik})$  is a parametric vector in the objective function;  $m$  ( $1, 2, 3, \dots, M$ ) is the number of input-output pairs. Essentially, the training process is to solve the multiobjective optimization problem with linear constraints, and the best parametric vector  $p(\beta_{jk}, \theta_k, \delta_{ik})$  can be obtained as a result of the training process, which is usually executed with certain optimization tool box of MATLAB software.

**4.2. Driving Anger Detection Model Based on BRB.** Considering different structures and sources of input data for driving anger detection model, a double-layered BRB (DL\_BRB) system is proposed in this study, as shown in Figure 5. The first layer consists of human (driver) status inference subsystem, vehicle status inference subsystem, and environment status inference subsystem. The second layer is to determine the driver's anger state based on the three subsystem's inference results, which are used as the inputs of the second layer. For the subsystem1, there are three human characteristics ( $x_1, x_2, x_3$ ) used as the inputs, while three vehicle characteristics ( $x_4, x_5, x_6$ ) and two environmental characteristics ( $x_7, x_8$ ) are used as the inputs for subsystem 2

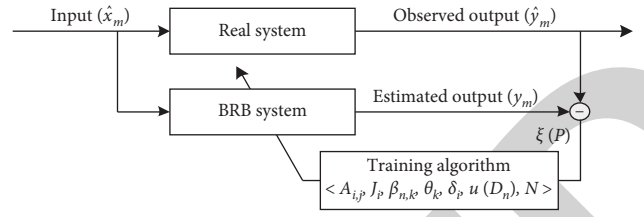


FIGURE 4: The training process of BRB inference system.

and subsystem 3, respectively. Especially,  $u_1, u_2, u_3$  are intermediate variables to denote driver status, vehicle status, and environment status, respectively, which are used to inference the driver's anger states ( $y$ ). The definition of all input variables ( $x_1, x_2, \dots, x_8$ ) and output variable ( $y$ ) is listed in Table 4. For example,  $x_1, x_2, x_3$  represent the driver's gender, age, and temperament, respectively.

As indicated in Figure 5, the input variables of the proposed model were comprised of three types of characteristics including human, vehicle, and environments, while the output variable of the proposed model is driving anger state ( $y$ ). Moreover, the human characteristics consisted of gender ( $x_1$ ), age ( $x_2$ ), and temperament ( $x_3$ ), the vehicle characteristics consisted of yaw angular acceleration ( $x_4$ ), longitudinal acceleration ( $x_5$ ), and lane departure ( $x_6$ ), and the environment characteristics consisted of traffic flow state ( $x_7$ ) and surrounding traffic behaviors ( $x_8$ ), shown in Table 4.

According to the distribution trait of those human, vehicle, and environment characteristics data, the interpretation or the reference values for every input variable with different grades were determined, as shown in Table 5. For instance, temperament, denoted by  $x_3$ , can be classified in to 4 different grades including melancholic, phlegmatic, sanguineous, and choleric, while the longitudinal acceleration, denoted by  $x_5$ , can be graded into 3 levels with references of  $1 \text{ m/s}^2$ ,  $1.5 \text{ m/s}^2$ , and  $2 \text{ m/s}^2$ , respectively. It is noted that the reference value for the different grades of yaw angular acceleration ( $x_4$ ), longitudinal acceleration ( $x_5$ ), and lane departure ( $x_6$ ) were determined according to their statistical data for the four driving anger states including neutral, low anger, moderate anger, and high anger. The relevant statics were studied in several previous research conducted by the authors in this study [20, 24, 29]. Moreover, traffic behavior around, denoted by  $x_8$ , was graded in to 3 different types, namely, normal, less severe, and severe, respectively. It is noted that waiting traffic lights belongs to normal traffic behavior around, and weaving/cutting in line or traffic jam/congestion belongs to less severe traffic behavior around, while jaywalking/motorcycle occupying road belongs to severe traffic behaviors around.

In this study, the initial belief rules of the proposed model were constructed based on the relevant expert knowledge and historical driving anger data with certain subjectivity. It is worth noting that, in terms of the second layer of the proposed model, all intermediate variables ( $u_1, u_2, u_3$ ) were graded into three levels, and then there exists 27 rules (33), shown in Table 6. Moreover, the final consequents were expressed in four types of driving anger states, namely,

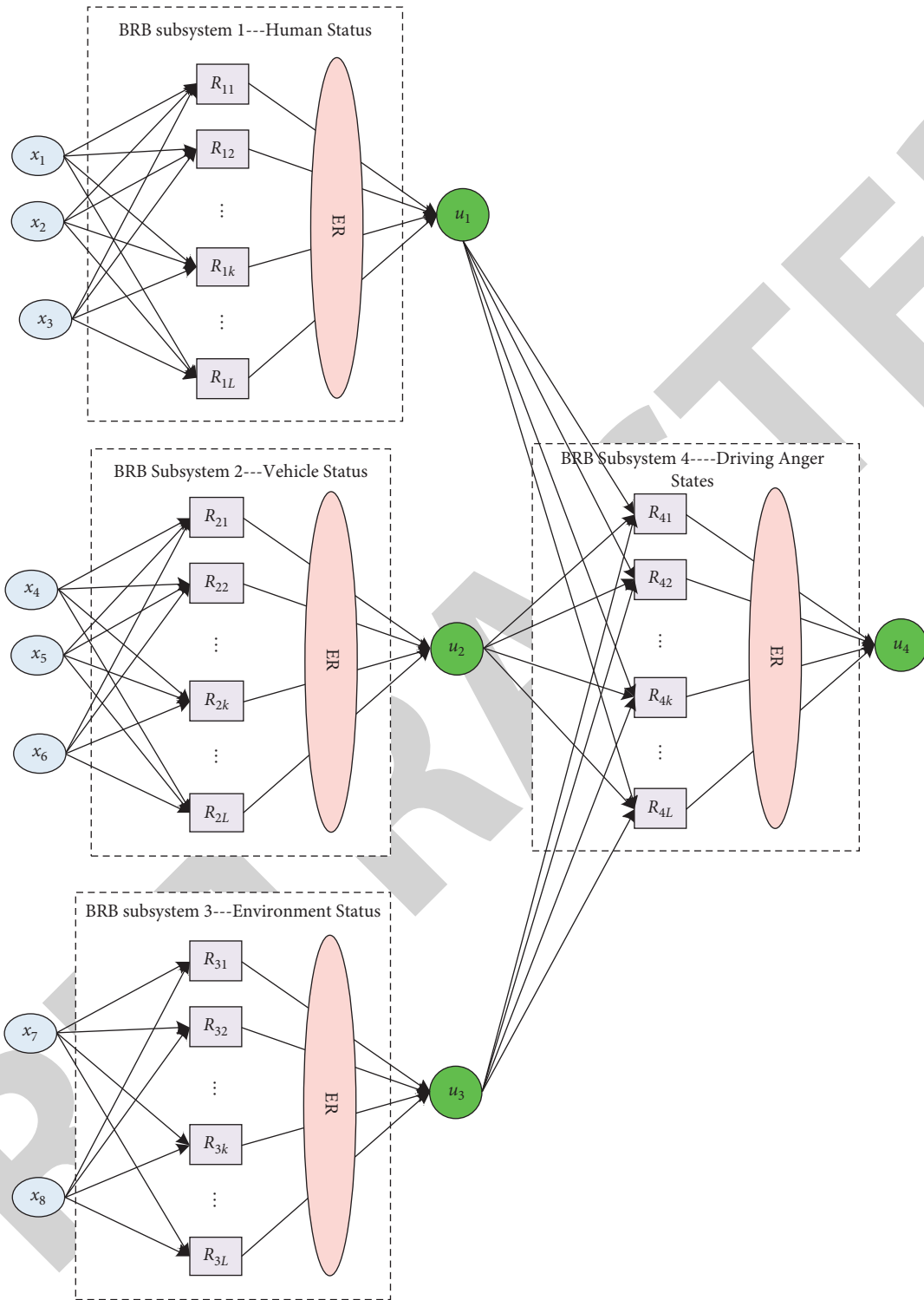


FIGURE 5: Schematic diagram of driving anger detection model based on DI\_BRB system.

neutral (none anger), low anger, moderate anger, and high anger, with a belief degree of  $\beta_{1k}, \beta_{2k}, \beta_{3k}$  and  $\beta_{4k}$  respectively.

As the initial belief rule base (BRB) is imprecise with certain subjectivity, the training of the initial BRB system was conducted according to equations (11) and (12). Then, 60% of samples were randomly selected from the original

2129 driving anger-related samples collected in the naturalistic experiments for training the initial BRB systems. The training processes were executed in optimization toolbox of Matlab, which mainly included computing the matching degree of the input ( $x_i$ ) to the reference value ( $A_i^k$ ), the activation weight of rule ( $R_k$ ), the belief degree of the  $j$ th consequent (output) in  $R_k$ , and minimizing the discrepancy

TABLE 4: Input variables and output variables of the proposed DI\_BRB system.

Antecedent attributes (human)		Antecedent attributes (environment)	
Gender	$x_1$	Traffic flow state	$x_7$
Age	$x_2$	Traffic behaviors around	$x_8$
Temperament	$x_3$		
Antecedent attributes (vehicle)		System output	
Yaw angular acceleration	$x_4$	Driving anger states	y
Longitudinal acceleration	$x_5$		
Lane departure	$x_6$		

TABLE 5: Graduation and reference value for the input and output variables.

Input	Grade	Reference value	Input	Grade	Reference value
$x_1$	1	Male-1	$x_6$	1	Small-0.350
	2	Female-2		2	Medium-0.562
$x_2$	1	Young-25		3	Large-0.725
	2	Average-35	$x_7$	Smooth	1
	3	Old-45		Less smooth	2
$x_3$	1	Melancholic-1		Obstructive	3
	2	Phlegmatic-2	Normal	1	
	3	Sanguineous-3	Less severe	2	
$x_4$	4	Choleric-4	$x_8$	Severe	3
	1	Low-1.75		Grade	Reference value
	2	Medium-2.00		Neutral	0
	3	High-2.50		Low anger	1
$x_5$	1	Low-1.0	Output	Moderate anger	2
	2	Medium-1.5		High anger	3
	3	High-2.0		Y	

TABLE 6: Initial belief rules of BRB subsystem 4.

Rule ID	Input U (intermediate variables)				Output y (belief degree)			
	Attribute weight	1	1	1	Ang_N	Ang_L	Ang_M	Ang_H
Rule weight	$u_1$	$u_2$	$u_3$	$\beta_{1k}$	$\beta_{2k}$	$\beta_{3k}$	$\beta_{4k}$	
1	1	1	1	1	1	0	0	0
2	1	1	1	2	0.8	0.1	0.1	0
3	1	1	1	3	0.7	0.1	0.1	0.1
4	1	1	2	1	0.6	0.2	0.1	0.1
5	1	1	2	2	0.5	0.3	0.1	0.1
6	1	1	2	3	0.5	0.2	0.2	0.1
7	1	1	3	1	0.4	0.3	0.2	0.1
8	1	1	3	2	0.4	0.4	0.1	0.1
9	1	1	3	3	0.3	0.4	0.2	0.1
10	1	2	1	1	0.3	0.3	0.3	0.1
11	1	2	1	2	0.3	0.4	0.2	0.1
12	1	2	1	3	0.3	0.3	0.2	0.2
13	1	2	2	1	0.2	0.4	0.2	0.2
14	1	2	2	2	0.2	0.2	0.3	0.3
15	1	2	2	3	0.2	0.2	0.2	0.4
16	1	2	3	1	0.1	0.5	0.2	0.2
17	1	2	3	2	0.1	0.4	0.3	0.2
18	1	2	3	3	0.1	0.2	0.4	0.3
19	1	3	1	1	0.1	0.2	0.4	0.3
20	1	3	1	2	0	0.1	0.5	0.4
21	1	3	1	3	0	0.2	0.4	0.4
22	1	3	2	1	0	0.2	0.3	0.5
23	1	3	2	2	0	0.2	0.2	0.6
24	1	3	2	3	0	0.1	0.2	0.7
25	1	3	3	1	0	0	0.2	0.8
26	1	3	3	2	0	0	0.1	0.9
27	1	3	3	3	0	0	0	1

between the observed output and the estimated output calculated by the proposed model. After finishing the training process, the optimal rule base of BRB subsystem 4 was obtained, as shown in Table 7.

#### 4.3. Test Results

**4.3.1. Evaluation Criteria.** To verify the identification accuracy of the DI\_BRB model proposed in this study, an evaluation method of receiver operating characteristic (ROC) curve analysis was implemented. The ROC method was widely used to evaluate classification performance of a classifier in discrimination fields like medical diagnosis, military monitoring, and human decision-making [30]. For a specific discrimination threshold, namely, cut-off point, it can differentiate positive samples from negative ones with a certain true positive rate (TPR) and false positive rate (FPR). When depicting the cut-off point in a coordinate system with horizontal and vertical ordinate, represented by TPR and FPR, respectively, a complete ROC curve will be formed by connecting all the possible cut-off points with a broken line in the coordinate system. Moreover, the greater the area under the ROC curve (AUC) is, the higher the identification accuracy of the classifier is. Except for TPR, FPR, and AUC, other statistical indexes such as positive predictive accuracy (PPA), *F1* and *Acc* are also extensively applied when evaluating a classifier's classification performance. Those statistical indexes are calculated as follows:

$$\begin{aligned}
 TPR &= \frac{TP}{TP + FN}, \\
 FPR &= \frac{FP}{TN + FP}, \\
 PPA &= \frac{TP}{TP + FP}, \\
 F1 &= \frac{2 \times TPR \times PPA}{TPR + PPA}, \\
 Acc &= \frac{TP + TN}{TP + FN + TN + FP},
 \end{aligned} \tag{13}$$

where *TP* is the number of positive samples, which were correctly identified as positive ones; *FP* is the number of negative samples, which are falsely identified as positive ones; *TN* is the number of negative samples, which are correctly identified as negative ones; *FN* is the number of positive samples, which are falsely identified as negative ones. Here, driving anger samples with a specific intensity are assumed to be positive, while the other driving anger samples are assumed to be negative. Note that *PPA* demonstrates the correct identification probability of the samples, which are identified as positive ones.

**4.3.2. Identification Performance.** The identification performance of the proposed model (DI\_BRB) for four driving anger states with different intensity was evaluated using a test set accounting for 40% of the total samples, which was

comprised of 374 neutral (none anger) samples, 218 low anger samples, 172 moderate anger samples, and 88 high anger samples. Further, in order to verify the superiority of the proposed model, we compared it with other 5 widely used classifiers like C4.5, NBC, SVM, kNN, and BPNN in terms of AUC of the corresponding ROC curves for the four driving anger states, illustrated in Figure 6.

As indicated in Figure 6, the identification performance of DI\_BRB model proposed in this study outperforms C4.5, NBC, KNN, SVM, and BPNN, as the proposed model has an average AUC of 0.8632, which is the biggest among them, according to the ROC curves for all driving anger states. Specially, the performance of the proposed model is evidently higher than the other five models when identifying neutral (none anger) state ( $p = 0.074$ ) and high anger state ( $p = 0.069$ ) in terms of AUC, yet the superiority is small and not significant when the significance level is set to be 0.05. In addition, when identifying the low anger state, the differences of AUC between the proposed model and other five models are much smaller than those of the other three driving anger states. Moreover, the average of AUC of all six identification models is the smallest for low anger state. The possible reason is that the changes of human, vehicle, and environmental characteristics under low anger state are evidently smaller than those of other driving anger states.

Furthermore, in order to evaluate the effectiveness of the proposed model, the five-fold cross-validation method was used to test the six classification models, with the results being indicated in Figure 7. We can see that the proposed DI\_BRB model achieves better identification performances in terms of PPA, TPR, and *F1* when compared with C4.5, BPNN, NBC, SVM, and kNN. Especially, the identification performances of the proposed model are remarkably ( $p = 0.038 < 0.05$ ) better than those of the other five models in terms of the three criteria for high anger state. However, for FPR, another important evaluation criterion, the proposed model does not perform the best, as the model of C4.5 has the lowest value of FPR when identifying all driving anger states.

In summary, the average of the aforementioned criteria and the total accuracy for identifying all driving anger states using the proposed model (DI\_BRB) and the other five widely implemented models were statistically analyzed, with the results being shown in Table 8. Then, we can see that the proposed model outperforms the other five models in most aspects including *TPR* (0.8433), *PPA* (0.9187), *F1* (0.8793), and *Acc* (0.8367). Though the model of C4.5 is superior to the proposed model in terms of FPR, the mean value of FPR of the proposed model is only 0.0792, ranking the second among these models, which is acceptable for detecting driving anger states in practice. Additionally, in order to verify the efficiency of the proposed model, the run speed of these models was compared in MATLAB environment for the same test data. Here, we let the runtime of BPNN model as unit one, and the runtime ratio of the other five models was computed, with the results being shown in the last row of Table 8. It is shown that the runtime ratio of the proposed model ranks the second, just behind C4.5, among the six models, indicating that the computation ability of the

TABLE 7: The trained belief rule base of BRB subsystem 4.

Rule ID	Input U (intermediate variable) attribute weight			Output y (belief degree)				
	Rule weight	$u_1$	$u_2$	$u_3$	Ang_N	Ang_L	Ang_M	Ang_H
1	0.9745	1	1	1	0.9925	0.0047	0.0025	0.0003
2	0.8742	1	1	2	0.7903	0.2084	0.0013	0.0000
3	0.8063	1	1	3	0.6900	0.2134	0.0933	0.0033
4	0.7551	1	2	1	0.6114	0.2981	0.0804	0.0001
5	0.7264	1	2	2	0.5093	0.2906	0.2001	0.0000
6	0.5552	1	2	3	0.4987	0.2047	0.1976	0.0990
7	0.4327	1	3	1	0.4128	0.3896	0.1072	0.0904
8	0.1685	1	3	2	0.3884	0.3125	0.1976	0.1015
9	0.3745	1	3	3	0.2124	0.3842	0.3142	0.0892
10	0.7264	2	1	1	0.4137	0.2817	0.3045	0.0001
11	0.5572	2	1	2	0.3125	0.3032	0.3019	0.0824
12	0.4613	2	1	3	0.2826	0.3216	0.2105	0.1853
13	0.4482	2	2	1	0.2256	0.4125	0.2753	0.0866
14	0.6164	2	2	2	0.1196	0.1316	0.3763	0.3725
15	0.7735	2	2	3	0.0015	0.1147	0.2872	0.5966
16	0.6981	2	3	1	0.0023	0.1873	0.4316	0.3788
17	0.7762	2	3	2	0.0046	0.0792	0.3247	0.5915
18	0.8614	2	3	3	0.0032	0.0041	0.2879	0.7048
19	0.4867	3	1	1	0.1145	0.2236	0.2974	0.3645
20	0.7786	3	1	2	0.0963	0.0842	0.2771	0.5424
21	0.6582	3	1	3	0.0074	0.1854	0.3882	0.4190
22	0.6963	3	2	1	0.0043	0.2352	0.2744	0.4861
23	0.4345	3	2	2	0.0172	0.2163	0.1882	0.5783
24	0.6625	3	2	3	0.0094	0.1125	0.1859	0.6922
25	0.5936	3	3	1	0.0015	0.0083	0.2305	0.7597
26	0.7652	3	3	2	0.0019	0.0026	0.1348	0.8607
27	0.9884	3	3	3	0.0003	0.0012	0.0043	0.9942

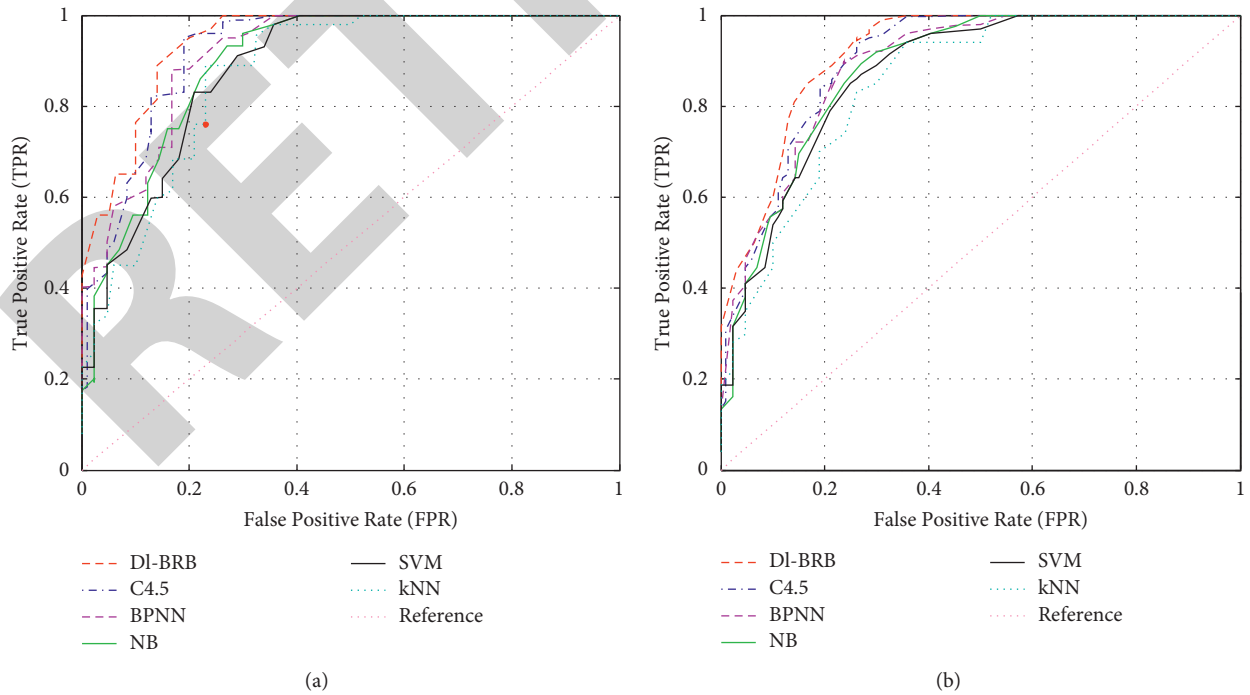


FIGURE 6: Continued.

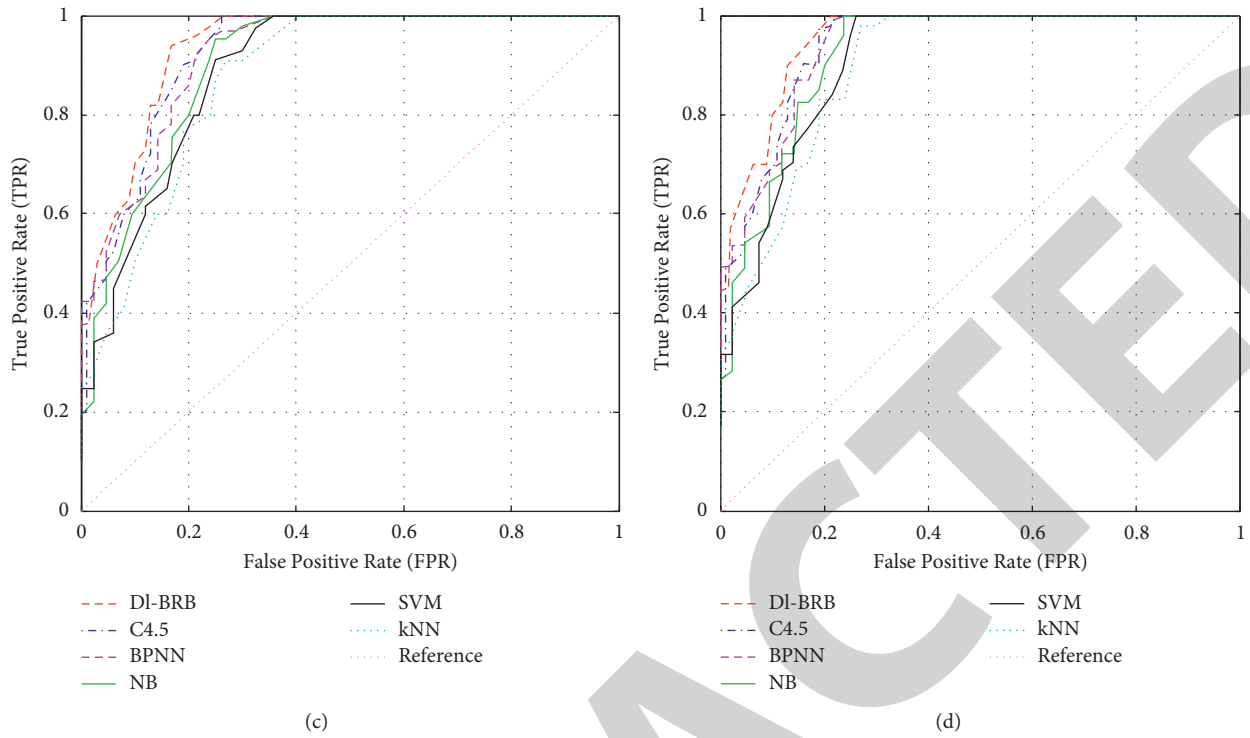


FIGURE 6: Comparison of ROC curves between DI\_BRB and other five widely used models when identifying four driving anger states. (a) Neutral. (b) Low anger. (c) Moderate anger. (d) High anger.

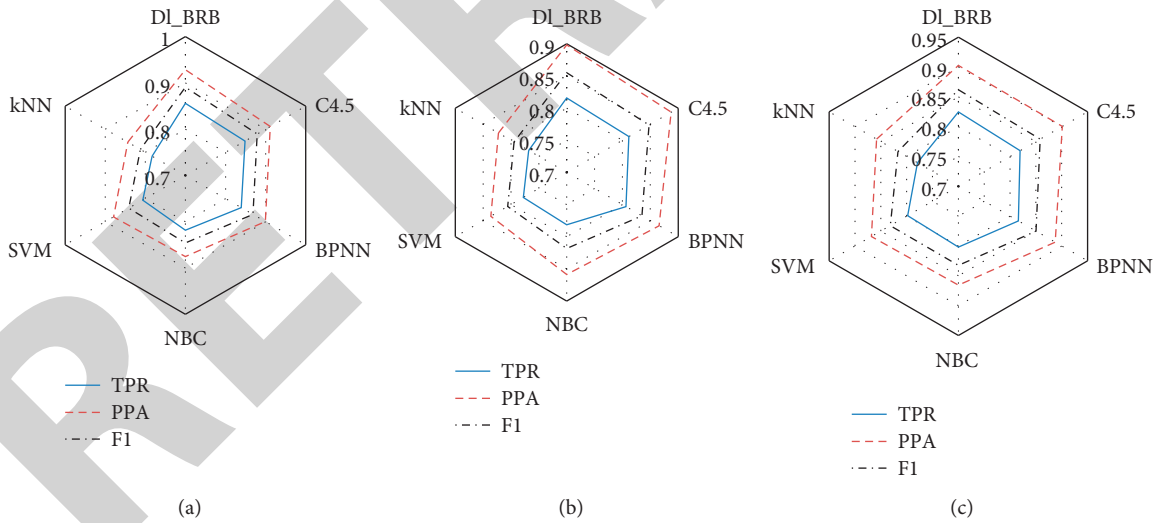


FIGURE 7: Continued.



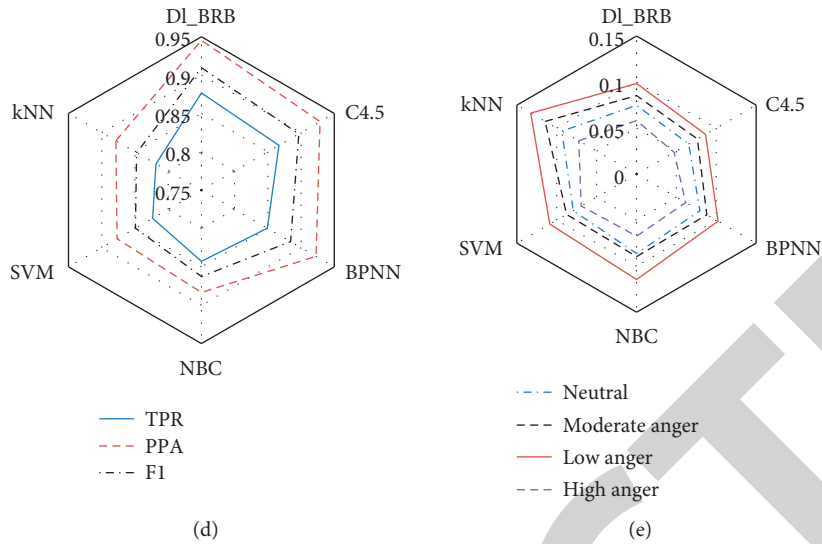


FIGURE 7: The identification performances of the proposed model and other five models. (a) TPR, PPA, and F1 for neutral state. (b) TPR, PPA, and F1 for low anger. (c) TPR, PPA, and F1 for moderate anger. (d) TPR, PPA, and F1 for high anger. (e) FPR of the four driving anger states.

TABLE 8: The average identification performances of the six models for four driving anger states.

	DL_BRB	C4.5	BPNN	NBC	SVM	kNN
TPR	0.8433	0.8366	0.8277	0.8111	0.8015	0.7872
PPA	0.9187	0.9072	0.8938	0.8708	0.8651	0.8509
F1	0.8793	0.8704	0.8594	0.8398	0.8321	0.8178
FPR	0.0792	0.0689	0.0830	0.0894	0.0862	0.1033
Acc	0.8367	0.8252	0.8215	0.8014	0.7792	0.7625
Runtime ratio	0.8824	0.8715	1.0000	0.8965	0.9845	0.9136

proposed model is enough for detecting driving anger in real time. In conclusion, the proposed DL\_BRB model is effective and efficient enough for detecting different driving anger states in practice.

## 5. Discussion and Conclusions

There are two contributions made in this study to the traffic psychology and driving behavior fields especially angry driving. The first one is that a novel driving anger elicitation approach was proposed to collect the relevant human (driver), vehicle, and environment characteristics in driving anger state. The second one is to propose an effective model based on those heterogeneous characteristics to identify different driving anger states with different intensity. Then, in practice, some soft interferences like releasing soft or relaxed music or comfortable human-machine interaction through conversation can be executed when low or moderate anger state is detected, while some hard interferences like controlling steering wheel or brake or acceleration pedal is taken over by human-machine codriving system when high anger state is detected in intelligent connected vehicle in the near future.

First, in order to induce real anger emotion during driving, a particularly busy route was selected as the test

route for the naturalistic experiments. On the test route, every participant inevitably came across sorts of incentive situations like jaywalking, motorcycle occupying road, weaving or cutting in line, traffic jam or congestion, and traffic light waiting, especially in morning or evening rush hours. The study results indicate that the hit rate of anger under the incentive situations reaches 74.25%, significantly higher than that under no incentive situations, and the highest hit rate (82.03%) of anger happened under the stimulus of jaywalking/motorcycle occupying road. Hence, some targeted and detailed countermeasures can be taken in practice. For example, once those kinds of uncivil (illegal) traffic behaviors are captured by cameras of electrical police, the personal information of pedestrian or the motorcycle rider like their portrait or the ID card number will be exposed on the big screen located on the intersection or roadside based on image identification and big data technology. Moreover, they might be fined if there is a real traffic police on the scene, as well as an increased insurance next year if they have driving license. In addition, every traffic participant should be encouraged to report those uncivil (illegal) traffic behaviors through some official APPs designed by traffic management authorities. Therefore, the novel anger elicitation approach proposed in this study is effective based on the induction of those incentive situations

in real environments and the stimuli of extra paid for finishing the whole test ahead of the reference time. Interestingly, there was inconsistency with previous findings, in which the most driving anger provoking event was discourtesy behavior from surrounding traffic participants in western countries like USA [31] and France [32]. The possible reason for that might be Chinese drivers often travelled with strong mixed traffic modes including motor and nonmotor vehicles and pedestrians competing with each other on the limited road lanes in a rapid pace of life nowadays. Additionally, different traffic rules, life style, culture background, and safety awareness also exerted an effect on the differences of driving anger provoking situations. Moreover, in China, Feng et al. [33] found that the most driving anger provoking situation for professional drivers was traffic obstruction; especially the road works sign was not set when the road was under reconstruction, showing a little disagreement with this study, in which Jaywalking/motorbike occupying road was the most driving anger provoking for private car drivers. Therefore, more target countermeasures should be put forward on monitoring and regulation of uncivilized or illegal travelling behaviors for driving training or traffic management authorities in China.

Secondly, according to previous research conducted by the authors in this study, the human characteristics like gender, age, and temperament, the environmental characteristics like traffic behaviors around and traffic flow state, and the vehicular characteristics like yaw angular acceleration, longitudinal acceleration, and lane departure have close relationship with driving anger [26, 29]. Then, those human, vehicular, and environmental characteristics were selected as the inputs for the establishing driving anger detection model. In this study, a data-driven model based upon belief rule base (BRB), which can continually utilize the accumulated filed data collected during the naturalistic experiments, is proposed to detect driving anger states with different intensity. The BRB based model has much advantage in processing not only objective information like vehicular characteristics data but also subjective information such as human and environmental characteristics data [27]. Moreover, a double-layered BRB (DI\_BRB) system including four BRB subsystems for inferring human, vehicular, and environmental state was constructed to reduce the number of rules to be estimated at the same time for improving the timely response. The study results show that the average identification accuracy (*Acc*) of the proposed model for detecting all driving anger states including neutral, low anger, moderate anger, and high anger reaches 83.67%, which is 1.15%, 1.52%, 3.53%, 5.75%, and 7.42% higher than C4.5, BPNN, NBC, SVM, and kNN, respectively. Besides, the proposed model is superior to those models in terms of other evaluation criteria such as TPR, PPA, and F1, indicating that the proposed model is effective to detect driving anger states in practice. However, the evaluation criterion of false positive rate (FPR) of the proposed model does not have superiority when compared with C4.5, a widely used decision tree algorithm. The reason for this is that maybe the inputs of the proposed model do not include any physiological

characteristics, which can precisely reflect the driver's mental fluctuation [34, 35]. Correspondingly, more physiological characteristics should be taken into consideration when constructing a driving anger detection model. In addition, the run time of the proposed model does not perform best amongst these widely used classification models, which means that the input variables set needs more optimization by feature selection algorithms to reduce the proposed model's computation complexity.

Nevertheless, there still exist several obvious limitations in this study. Firstly, considering easy recruiting and significant anger induction effect, much more male drivers were recruited as the participants for the naturalistic experiments conducted in this study. Then, much more female drivers should be recruited to improve the universality of the proposed model afterwards. Additionally, female drivers frequently get angrier than males when confronting traffic block. Nonetheless, they are inclined to make more adaptive adjustments instead of aggression behaviors when they get angry [22]. Thus, more traffic situations like traffic congestion, red light waiting, and jaywalking can be used to trigger anger for female drivers. And the measurements including physiological or facial expression characteristics instead of physical behaviors and verbal aggressions can be utilized to demonstrate anger expression differences between female drivers and male drivers. Secondly, the naturalistic experiments for this study were conducted in Wuhan, a typical central metropolis in China. Nevertheless, there exist some geography variances of pace of life, traffic safety awareness, or quality between Wuhan and other typical western cities (e.g., Chengdu), northern cities (e.g., Beijing), eastern cities (e.g., Shanghai), and southern cities (e.g., Guangzhou). Therefore, the succeeding experiments should be added in those typical cities due to the geography differences, which can result in the differences of anger induction factors. Thirdly, with the rapid development of wearable devices and intelligent connected vehicles, some physiological indicators like heart rate and respiration rate can be collected to detect driving anger states more accurately.

### Data Availability

The data are generated from experiments and can be made available by the corresponding author upon request.

### Conflicts of Interest

The authors declare that they have no conflicts of interest.

### Acknowledgments

The present study was supported by projects funded by National Nature Science Foundation of China (52062014), Jiangxi University Humanities and Social Science Research Project (GL19204), Natural Science Foundation of Jiangxi Province (20212BAB202010), and Fund of Science and Technology program of Education Department of Jiang Xi Province (GJJ180359) and also supported by National

Nature Science Foundation of China (52162049 and 52062015) and Jiangxi Provincial Major Science and Technology Project-5G Research Project (20212ABC03A07).

## References

- [1] C.-Z. Wu and H. Lei, "Review on the study of motorists' driving anger," *China Safety Science Journal*, vol. 20, no. 7, pp. 3–8, 2010.
- [2] Nhtsa, *Traffic Safety Facts 2007: A Compilation of Motor Vehicle Crash Data from the Fatality Analysis Reporting System and the General Estimates System*, National Highway Traffic Safety Administration, Washington DC, USA, 2007.
- [3] S.-Y. Lv, "The Ministry of Public Security will crack down on "road rage" violations such as forced lane changes and overtaking," 2016, <http://finance.chinanews.com/auto/2015/05-11/7266660.shtml>.
- [4] E. R. Dahlen, R. C. Martin, K. Ragan, and M. M. Kuhlman, "Driving anger, sensation seeking, impulsiveness, and boredom proneness in the prediction of unsafe driving," *Accident Analysis & Prevention*, vol. 37, no. 2, pp. 341–348, 2005.
- [5] T. Lajunen and D. Parker, "Are aggressive people aggressive drivers? a study of the relationship between self-reported general aggressiveness, driver anger and aggressive driving," *Accident Analysis & Prevention*, vol. 33, no. 2, pp. 243–255, 2001.
- [6] L. Kessous, G. Castellano, and G. Caridakis, "Multimodal emotion recognition in speech-based interaction using facial expression, body gesture and acoustic analysis," *Journal on Multimodal User Interfaces*, vol. 3, no. 1, pp. 33–48, 2010.
- [7] P. N. Juslin and J. A. Sloboda, *Handbook of Music and Emotion: Theory, Research, Applications*, Oxford University Press, New York, NY, U.S.A., 2010.
- [8] H. Lei, X.-P. Yan, and C.-Z. Wu, "The characteristic of vehicle speed under angry driving in China," in *Proceedings of the 2nd International Conference on Transportation Information and Safety*, pp. 1542–1547, Wuhan, China, June 2013.
- [9] E. Roidl, F. W. Siebert, M. Oehl, and R. Höger, "Introducing a multivariate model for predicting driving performance: the role of driving anger and personal characteristics," *Journal of Safety Research*, vol. 47, pp. 47–56, 2013.
- [10] R. Abdu, D. Shinar, and N. Meiran, "Situational (state) anger and driving," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 15, no. 5, pp. 575–580, 2012.
- [11] H. Cai and Y. Lin, "Modeling of operators' emotion and task performance in a virtual driving environment," *International Journal of Human-Computer Studies*, vol. 69, no. 9, pp. 571–586, 2011.
- [12] M. Danaf, M. Abou-Zeid, and I. Kaysi, "Modeling anger and aggressive driving behavior in a dynamic choice-latent variable model," *Accident Analysis & Prevention*, vol. 75, pp. 105–118, 2015.
- [13] K. Oeltze and C. Schießl, "Benefits and challenges of multi-driver simulator studies," *IET Intelligent Transport Systems*, vol. 9, no. 6, pp. 618–625, 2015.
- [14] A. J. Flidlund and E. Z. Izard, *Electromyographic Studies of Facial Expressions of Emotions and Patterns of Emotions*, Guilford Press, New York, NY, U.S.A., 1983.
- [15] J. Wang and Y. Gong, "Normalizing multi-subject variation for drivers' emotion recognition," in *Proceedings of the IEEE International Conference on Multimedia and Expo*, vol. 3, pp. 354–357, New York, NY, U.S.A., July 2009.
- [16] C. D. Katsis, Y. Goletsis, G. Rigas, and D. I. Fotiadis, "A wearable system for the affective monitoring of car racing drivers during simulated conditions," *Transportation Research Part C: Emerging Technologies*, vol. 19, no. 3, pp. 541–551, 2011.
- [17] P. Wan, C. Wu, Y. Lin, and X. Ma, "On-road experimental study on driving anger identification model based on physiological features by ROC curve analysis," *IET Intelligent Transport Systems*, vol. 11, no. 5, pp. 290–298, 2017.
- [18] L. Malta, C. Miyajima, N. Kitaoka, and K. Takeda, "Analysis of real-world driver's frustration," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 1, pp. 109–118, 2011.
- [19] A. Lanatà, G. Valenza, A. Greco et al., "How the autonomic nervous system and driving style change with incremental stressing conditions during simulated driving," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, pp. 1505–1517, 2015.
- [20] P. Wan, C.-Z. Wu, Y.-Z. Lin, and X.-F. Ma, "Driving anger detection based on multivariate time series features of driving behaviors," *Journal of Jilin University (Engineering and Technology Edition)*, vol. 47, no. 5, pp. 1426–1435, 2017.
- [21] A. Muñoz, M. Domingo, M. Trinidad, and J. Delgado, "Integration of body sensor networks and vehicular ad-hoc networks for traffic safety," *Sensors*, vol. 16, no. 107, pp. 1–29, 2016.
- [22] G. I. Beatriz, G. F. Jose Antonio, and L. M. Angeles, "Driving anger and traffic violations: gender differences," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 15, pp. 404–412, 2012.
- [23] Z. Chen, C. Wu, M. Zhong, N. Lyu, and Z. Huang, "Identification of common features of vehicle motion under drowsy/distracted driving: a case study in Wuhan, China," *Accident Analysis & Prevention*, vol. 81, pp. 251–259, 2015.
- [24] P. Wan, C.-Z. Wu, and X.-F. Ma, "Discriminating threshold of driving anger intensity based on driving behavior features by ROC curve analysis," *Journal of Jilin University (Engineering and Technology Edition)*, vol. 50, no. 1, pp. 121–131, 2020.
- [25] J. Liu, P. Tian, J. Rong, and F.-T. Ren, "Initial Research on relationship between Driver's temperament and travel speed," *Journal of Beijing University of Technology*, vol. 32, no. 1, pp. 27–32, 2006.
- [26] P. Wan, L.-X. Yan, and S. Lu, "Sensitivity analysis of incentive situations and personal attributes to driving anger based on mnl model: a naturalistic experimental study," *Technical Gazette*, vol. 28, no. 5, pp. 1654–1666, 2021.
- [27] J.-B. Yang, J. Liu, J. Wang, H.-S. Sii, and H.-W. Wang, "Belief rule-base inference methodology using the evidential reasoning approach-RIMER," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 36, no. 2, pp. 266–285, 2006.
- [28] H. Zhu, J.-B. Zhao, Y. Xu, and L.-M. Du, "Interval-valued belief rule inference methodology based on evidential reasoning-IRIMER," *International Journal of Information Technology and Decision Making*, vol. 15, no. 6, pp. 1345–1366, 2016.
- [29] P. Wan, C.-Z. Wu, Y.-Z. Lin, and X.-F. Ma, "Driving anger states detection based on incremental association markov blanket and least square support vector machine," *Discrete Dynamics in Nature and Society*, vol. 2019, Article ID 2745381, 17 pages, 2019.
- [30] T. Bechtel, L. Capineri, C. Windsor, and M. Inagaki, "Comparison of ROC curves for landmine detection by holographic radar with ROC data from other methods," in *Proceedings of the 8th International Workshop on Advanced Ground Penetrating Radar*, pp. 1–4, Florence, Italy, July 2015.
- [31] J. L. Deffenbacher, A. N. Stephens, and M. J. M. Sullman, "Driving anger as a psychological construct: twenty years of

## Research Article

# Evaluation of Regional Ecological Efficiency and Intelligent Decision Support for Sustainable Development Based on Environmental Big Data

Lelai Shi <sup>1</sup>, Jing Zhou <sup>1</sup>, and Xingchi Zhou<sup>2</sup>

<sup>1</sup>Center for Industrial Economic Studies, School of Economics, Wuhan Textile University, Wuhan 430200, China

<sup>2</sup>School of Management, Wuhan Textile University, Wuhan 430200, China

Correspondence should be addressed to Jing Zhou; [ccedit@hotmail.com](mailto:ccedit@hotmail.com)

Received 2 November 2021; Revised 18 December 2021; Accepted 24 December 2021; Published 27 January 2022

Academic Editor: Daqing Gong

Copyright © 2022 Lelai Shi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

To promote urbanization in the next stage, it is of great significance to explore the ecological efficiency of green ecological regions and advance the sustainable development of a social economic system. However, spatial heterogeneity has not been fully considered in the existing evaluation models or methods for regional ecological efficiency (REE), and the corresponding decisions on sustainable development are not the optimal solutions. To solve the problems, this paper explores the evaluation of REE and intelligent decision support for sustainable development by analyzing environmental big data. Firstly, the spatiotemporal evolution of REE was examined based on environmental big data to clarify the spatial layout of REE and the sources of the spatial differences. Next, a multiobjective optimal decision-making model was established for the sustainable development of a regional ecosystem, and the solving method was presented for the model. The proposed model was proved valid through experiments.

## 1. Introduction

Urban development faces several problems, namely, the development model is extensive, the newly developed areas are underpopulated, and the resources are in short supply. With the development of the industry, these problems have intensified the contradiction between population, environment, and resources [1–5]. Rather than pursuing the single goal of economic interest, the sustainable development of the urbanization system aims to organize production for the composite goal of resources/energy, environment, economy, and society.

China is vigorously implementing a strategy of regional ecosystem development and environmental protection, which will continuously enhance the coordination between the regional environment, regional economy, and regional society [6–12]. Against this backdrop, in order to promote urbanization in the next stage, it is of great significance to explore the ecological efficiency of green ecological regions and advance the sustainable development of the social economic system [13–21].

Environmental pollution is a potential bottleneck of stable economic growth. As an effective measure of sustainable development, ecological efficiency can fully reflect the actual level of coordinated development between economy and environment. Liu and Sun [22] constructed a data envelopment analysis (DEA) model with environmental pollution and resource consumption as inputs and total economy as the output. Ratner [23] described the monitoring of ecological and economic efficiencies of the activities in the regional economic system, compared the applicability of several methods to the dynamic DEA model, and attempted to generate time series based on environmental and economic efficiency points. To evaluate the performance of the urban system, Giordano et al. [24] divided the urban system into traffic subsystem, built environment subsystem, and social economy subsystem, evaluated the efficiency of each subsystem by Takagi–Sugeno model, and recursively derived the overall efficiency of the entire system. Hoang and Alauddin [25] designed an input-oriented DEA framework, which allows the measurement



and decomposition of economic, environmental, and ecological efficiencies in agricultural production of different countries.

On the sustainable development of the urban ecological system, Wang and Li [26] set up an overall framework for the sustainable urban spatial development model based on smart cities. The framework, involving such four dimensions as the monitoring, collection, interconnection, and sharing of urban data, provides an important guide for the spatial planning and construction of smart cities. Ma et al. [27] analyzed the evolution of an urban ecosystem and its dissipation structure and constructed an evaluation index system (EIS) for the sustainable development capability of the urban ecosystem.

The existing studies provide some useful references. However, the evaluation models and methods for regional ecological efficiency fail to fully consider spatial heterogeneity, take account of the dynamic evolution of ecological efficiency distribution, and thoroughly discuss the key factors affecting the ecological efficiency identification and animal habitat changes in different regions and at different levels. Concerning sustainable development, the available decision-making models overlook the balance of relative development for the regional ecosystem in spatiotemporal evolution and cannot converge to the optimal decisions. To solve the problems, this paper explores the evaluation of REE and intelligent decision support for sustainable development based on environmental big data. Section 2 analyzes the spatiotemporal evolution of REE based on environmental big data, identifies the spatial layout and spatial difference of REEs, and discusses the key environmental factors. Section 3 establishes a multiobjective optimal decision-making model for the sustainable development of the regional ecosystem and presents the solving method for the model. Experimental results demonstrate the effectiveness of our model.

This research discusses the inherent linking mechanism of system sustainable development with environmental, economic, and social impacts, and optimizes the absolute performance of sustainability and relative development simultaneously.

## 2. Spatiotemporal Evolution Analysis

This paper identifies the sources of environmental data for the spatiotemporal evolution of REE. Table 1 lists the REE evaluation items based on big data. Several environmental types are enumerated, namely, building area, traffic area, water area, and landscape area, and the corresponding factors that affect ecological efficiency are presented in detail.

As mentioned before, this paper aims to clarify the spatial layout of REE and the sources of the spatial differences and realize the REE evaluation and make intelligent decisions for sustainable development. For this purpose, the spatial distribution of ecological efficiency was described intuitively based on environmental big data. Then, the degree of spatial differences of REE was investigated, and the sources of these differences were revealed. Finally, the dynamic evolution law of REE was explored through kernel density estimation (KDE) and Markov chain.

*2.1. Spatial Difference Analysis.* Through Dagum's decomposition of the Gini coefficient, this paper analyzes the spatial differences of global and local ecological efficiencies of the ecosystem in the study area. Dagum's decomposition method was selected for its capability of considering the overlap between subsamples, solving heterogeneous sources of REE spatial differences, and measuring how much overall regional differences are affected by the intraregional difference, inter-regional difference, and transvariation intensity.

Let  $b_{uv}$  and  $b_{gs}$  be the urban ecological efficiencies of the  $u$ -th and  $g$ -th regions, respectively; let  $\lambda$  be the mean ecological efficiency of all cities in a region,  $m$  be the total number of samples (cities),  $l$  be the number of regions, and  $m_u$  and  $m_g$  be the number of cities in the  $u$ -th and  $g$ -th regions, respectively. Then, the relative difference between all cities in a region in ecological efficiency can be measured by the regional overall Gini coefficient:

$$GN = \frac{\sum_{u=1}^l \sum_{g=1}^l \sum_{v=1}^{m_u} \sum_{s=1}^{m_g} |b_{uv} - b_{gs}|}{2m^2\lambda}. \quad (1)$$

To decompose the Gini coefficient by a regional subsystem, the first step is to sort all cities in a region by mean ecological efficiency as follows:

$$\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_g \leq \dots \leq \lambda_u \leq \dots \leq \lambda_l. \quad (2)$$

Ecological efficiency is the ratio of ecological outputs to ecological inputs. The outputs refer to the value of products and services provided by enterprises or economies, while the inputs refer to the resources and energies consumed by enterprises or economies, as well as the environmental load caused by enterprises or economies. The Gini coefficient of the  $u$ -th region can be calculated by

$$GN_u = \frac{(1/2\lambda_u) \sum_{v=1}^{m_u} \sum_{s=1}^{m_u} |b_{uv} - b_{us}|}{m_u^2}. \quad (3)$$

The intraregional contribution difference  $H_q$  can be calculated by

$$H_q = \sum_{u=1}^l GN_u t_u. \quad (4)$$

The inter-regional Gini coefficient  $GN_{ug}$  between the  $u$ -th and  $g$ -th regions can be calculated by

$$GN_{ug} = \frac{\sum_{v=1}^{m_u} \sum_{s=1}^{m_g} |b_{uv} - b_{gs}|}{m_u m_g (\lambda_u + \lambda_g)}. \quad (5)$$

The inter-regional contribution difference  $H_{mr}$  can be calculated by

$$H_{mr} = \sum_{u=2}^l \sum_{h=2}^{u-1} GN_{ug} (t_u e_h + t_h e_u) C_{ug}. \quad (6)$$

The transvariation intensity contribution  $H_d$  can be calculated by

TABLE 1: REE evaluation based on environmental big data.

Environmental type	Parameters	Influencing factors of ecological efficiency
Building area	Newly built area, energy consumption per unit area, and energy consumption structure	Energy-efficient design, building energy conservation and emission reduction, and green building material utilization
Traffic area	Passenger capacity per unit time	Public transit travel rate and green travel rate
Water area	Water supply and sewage treatment capacity	Water-saving equipment usage, water-saving indices, and wastewater and sewage reuse rate
Landscape area	Green area and greening rate	Vegetation coverage and regional carbon sequestration capacity

$$H_d = \sum_{u=2}^l \sum_{g=2}^{u-1} \text{GN}_{ug}(t_u e_g + t_g e_u)(1 - C_{ug}), \quad (7)$$

where  $e_i = m_i \lambda_i / (m \lambda)$  and  $t_j = m_j / m$ . Let  $c_{ug}$  be the ecological efficiency difference between regions;  $t_{ug}$  be the first moment of transvariation. Then, the relative influence  $C_{ug}$  of the unit ecological efficiency of the  $u$ -th and  $g$ -th regions can be calculated by

$$C_{ug} = \frac{c_{ug} - t_{ug}}{c_{ug} + t_{ug}}. \quad (8)$$

Let  $O_u$  and  $O_g$  be the cumulative density distribution functions of the  $u$ -th and  $g$ -th regions, respectively. Then,  $c_{ug}$  can be characterized by the mathematical expectation of the sum of the positive sample values for the difference between the ecological efficiencies,  $b_{uv}$  and  $b_{gs}$ , of the cities in the  $u$ -th and  $g$ -th regions:

$$c_{ug} = \int_0^{\infty} dO_u(b) \int_0^y (b-a) dO_g(a). \quad (9)$$

In addition,  $t_{ug}$  can be calculated by

$$t_{ug} = \int_0^{\infty} dO_g(b) \int_0^y (b-a) dO_u(a). \quad (10)$$

**2.2. Dynamic Evolution of Distribution.** The traditional KDE, a popular tool in spatial nonequilibrium analysis, is a nonparametric estimation approach. It can describe the form of each random variable with a continuous curve. In this way, virtually no statistical error will be incurred by the improper setting of ecological efficiencies in the region. Let  $M$  be the total number of cities in the region,  $A_i$  be the independent identically distributed observations,  $a^*$  be the mean value,  $\tau$  be the bandwidth, and  $\Gamma(\cdot)$  be the kernel function. Then, the density function  $\mu(a)$  of the random variable  $A$  can be calculated by

$$\mu(a) = \frac{1}{M\tau} \sum_{i=1}^M \Gamma\left\{\frac{(A_i - a^*)}{\tau}\right\}. \quad (11)$$

The greater the  $\tau$  value, the larger the neighbourhood of  $a$ . If  $\tau$  is too large, the kernel density function  $\Gamma(\cdot)$  will be too smooth. Then, some important features of the function will be smoothed out, causing a large deviation. Hence, the  $\tau$  value should be minimized if possible. Thus,  $\Gamma(\cdot)$  needs to meet the following condition:

$$\begin{cases} \lim_{a \rightarrow \infty} \Gamma(a) \cdot a = 0, \\ \Gamma(a) \geq 0, \quad \int_{-\infty}^{+\infty} \Gamma(a) da = 1, \\ \sup \Gamma(x) < +\infty, \quad \int_{-\infty}^{+\infty} \Gamma(a) da < +\infty. \end{cases} \quad (12)$$

The dynamic evolution features of REE distribution can be estimated by the Gaussian kernel function:

$$\Gamma(a) = \frac{1}{\sqrt{2\pi}} e^{-a^2/2}. \quad (13)$$

**2.3. Long-Term Transition Trend.** This paper adopts the Markov chain to describe the ecological efficiency of each city in a region, that is, sets up the transition probability matrix of the Markov chain. Let  $\{A(p), p \in \psi\}$  be the stochastic process of the discrete events corresponding to the Markov chain,  $N$  be the finite set of the solutions, and  $\psi$  be the set of indices for different stages of the stochastic process. The Markov chain highlights that, historical dynamic actions share the same features as future dynamic actions. In our problem, the identicalness of the features manifests as follows: the state of urban ecological efficiency  $A$  in year  $p+1$  directly bears on the probability for  $A$  to belong to type  $j$  in year  $p$ . Let  $GR_{ij}$  be the probability for urban ecological efficiency belonging to type  $i$  in year  $p$  to transfer to type  $j$  in year  $p+1$ ; that is, the maximum likelihood estimation of the transition is  $GR_{ij} = m_{ij}/m_i$ ; let  $m_{ij}$  be the number of cities shifting from type  $k$  in year  $p$  to type  $j$  in year  $p+1$  and  $m_i$  be the total number of cities belonging to type  $i$  in the study period. Then, we have

$$\begin{aligned} GR\{A_p = j | A_{p-1} = i, A_{p-2} = i_{p-2}, A_0 = i_0\} \\ = GR\{A_m = j | A_{m-1} = i\} = GR_{ij}. \end{aligned} \quad (14)$$

By dividing urban ecological efficiencies in the region into  $x$  states, it is possible to obtain an  $x \times x$  state transition probability matrix of the Markov chain.

### 3. Multiobjective Optimal Decision-Making Model

**3.1. Preliminaries.** Figure 1 summarizes the evaluation attributes for the sustainable development of the regional ecosystem. The evaluation system for the sustainable development of the regional ecosystem involves multiple



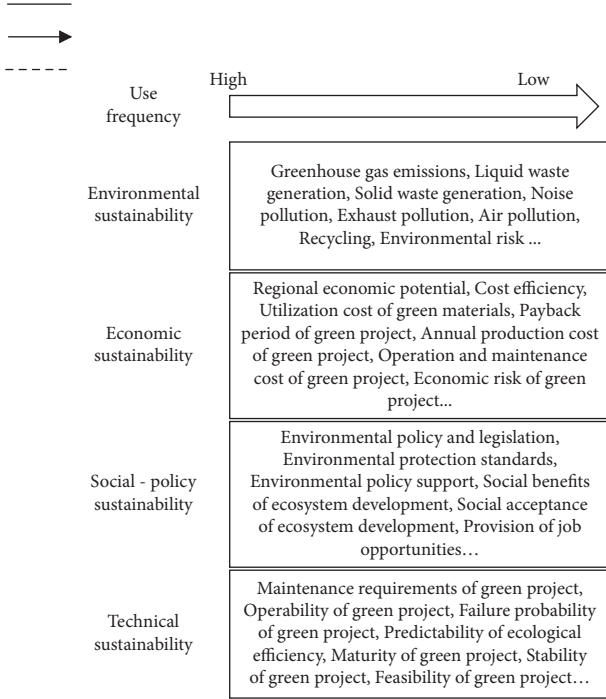


FIGURE 1: Evaluation attributes for the sustainable development of regional ecosystem.

dimensions, such as environmental sustainability, economic sustainability, social policy sustainability, and technical sustainability. The evaluation dimensions can all be quantified. For the sustainable development of the regional ecosystem, decision makers can also set up individualized evaluation systems based on objective needs and subjective judgement.

To clarify the incremental costs and benefits of regional ecosystem development, this paper focuses on the incremental costs incurred during the optimization of plan design and measures implementation for regional ecosystem development. The costs of these two stages are mainly invested continuously by green project investors, developers, and implementers. The incremental benefits will continue to appear after the optimization measures are implemented in the regional ecosystem. Figure 2 details the costs and benefits of the regional ecosystem in a development cycle.

Our problem is to sustainably improve the regional ecosystem, that is, combining and optimizing multiple optional measures for the ecosystem optimization to realize multiple goals under specific constraints, in reference to multiple sets of environmental big data in different periods, such that the optimized regional ecosystem is much more sustainable than the original system. This paper builds a weighted multiobjective optimization model to improve the sustainability of the regional ecosystem in multiple dimensions, relying on optimization measures. On this basis, the authors strived to make sustainable intelligent decisions.

**3.2. Model Construction and Solving.** Targeting the regional ecosystem, our model tries to improve the ecological efficiency of the ecosystem by adopting the improved

combination of optimization measures for sustainable development. Figure 3 provides the specific steps to improve sustainable development efficiency: characterization of the regional ecosystem, sustainable development evaluation of REE, sustainable development optimization of REE, and sustainable development improvement of REE.

The optimal decision making for the sustainable development of REE involves multiple steps and goals in four dimensions, including environmental sustainability, economic sustainability, social policy sustainability, and technical sustainability. The decision-making process is complicated by the diverse constraints on the different goals and the interference of uncertain decision-making factors. Hence, this paper presents a multiobjective optimal decision-making model (Figure 4) for the sustainable development of the regional ecosystem. The proposed model consists of two parts, namely, the characterization of the ecosystem, and construction and solving of sustainable development decision-making model and opens the path to improve the sustainable development of the regional ecosystem through spatiotemporal evolution.

The proposed multiobjective optimal decision-making model is targeted at regional ecosystems, whose ecological efficiencies obey different spatial distributions. The model was constructed through composite weighting of each subobjective, setting weighted objectives and sustainable objectives, and solving the model. Firstly, the interval analytic hierarchy process (IAHP) was adopted to assign composite weight to each objective, eliminating the need to consider the interaction between subobjectives. The composite weighting begins with the construction of an interval comparison matrix.

The subjective weight of each subobjective was quantified through expert judgement, which is uncertain to a certain extent. Therefore, the comparison matrix is composed of the following interval numbers:

$$F = \begin{bmatrix} [1, 1] & [f_{12}^K, f_{12}^V] & \cdots & [f_{1m}^K, f_{1m}^V] \\ [f_{21}^K, f_{21}^V] & [1, 1] & \cdots & [f_{2m}^K, f_{2m}^V] \\ \vdots & \vdots & \ddots & \vdots \\ [f_{m1}^K, f_{m1}^V] & [f_{m2}^K, f_{m2}^V] & \cdots & [1, 1] \end{bmatrix}, \quad (15)$$

where  $m$  is the number of optimization subobjectives;  $[f_{ij}^K, f_{ij}^V]$  is the relative importance of subobjective  $i$  to subobjective  $j$ , both of which are expressed as interval numbers. Next, the interval comparison matrix  $F$  was converted into the interval priority matrix  $S$ :

$$S = \begin{bmatrix} [0.5, 0.5] & [s_{12}^K, s_{12}^V] & \cdots & [s_{1m}^K, s_{1m}^V] \\ [s_{21}^K, s_{21}^V] & [0.5, 0.5] & \cdots & [s_{2m}^K, s_{2m}^V] \\ \vdots & \vdots & \ddots & \vdots \\ [s_{m1}^K, s_{m1}^V] & [s_{m2}^K, s_{m2}^V] & \cdots & [0.5, 0.5] \end{bmatrix}, \quad (16)$$

where

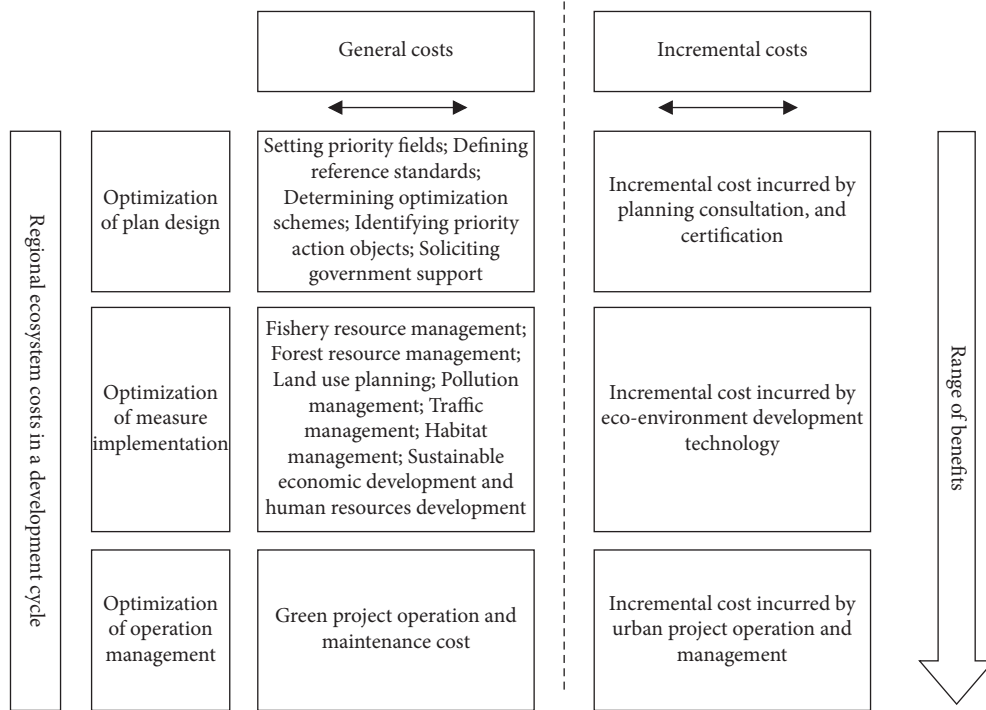


FIGURE 2: Costs and benefits of regional ecosystem in a development cycle.

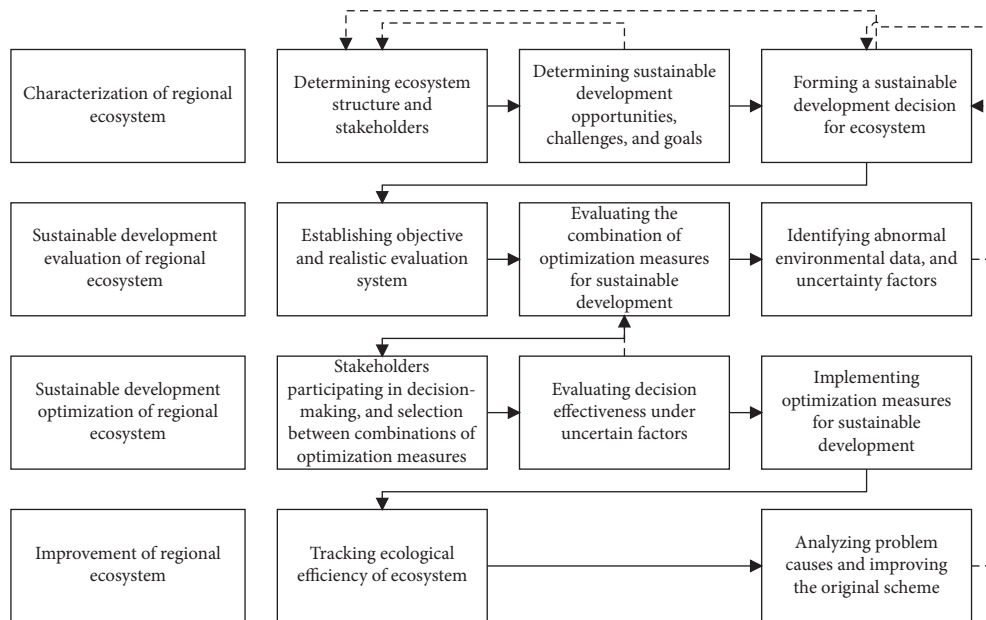


FIGURE 3: Improvement of sustainable development efficiency.

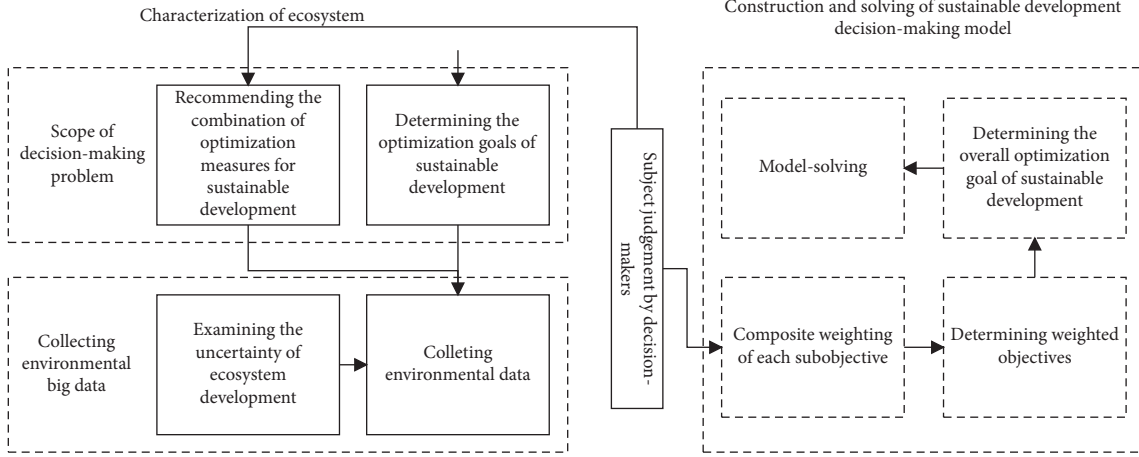


FIGURE 4: Multiobjective optimal decision-making model for the sustainable development of regional ecosystem.

$$\bar{s}_{ij} = \frac{f_{ij}^-}{f_{ij}^- + f_{ij}^+}, s_{ij}^+ = \frac{f_{ij}^+}{f_{ij}^- + f_{ij}^+}. \quad (17)$$

After that, the subjective weights were calculated. Let  $d_i^V$ ,  $d_i^K$ ,  $c_i^V$ , and  $c_i^K$  be the positive and negative biases in objective programming, respectively;  $[v_i^K, v_i^V]$  be the interval value range of the relative importance for each subobjective corresponding to the optimal solution. Then, the optimal solution of matrix  $S$  can be obtained by

$$\begin{aligned} \text{Min } \delta &= \sum_{i=1}^m (d_i^V + d_i^K + c_i^V + c_i^K), \\ \text{s.t. } v_i^K + \sum_{j=1, j \neq i}^m v_j^V &\geq 1, \\ v_i^U + \sum_{j=1, j \neq i}^m v_j^K &\leq 1, \\ 0 \leq v_i^K \leq v_i^V &\leq 1, \\ \left( \sum_{j=1}^m s_{ij}^K - m + 0.5 \right) v_i^K + \sum_{j=1, j \neq i}^m r_{ij}^K v_j^V - d_i^V + d_i^K &= 0, \\ \left( \sum_{j=1}^m s_{ij}^V - m + 0.5 \right) v_i^V + \sum_{j=1, j \neq i}^m s_{ij}^V v_j^K - c_i^V + c_i^K &= 0, \\ d_i^V, d_i^K, c_i^V, c_i^K &\geq 0. \end{aligned} \quad (18)$$

The subjective weight of each objective can be obtained by

$$q_i^e = \frac{(v_i^K + v_i^V)}{\sum_{i=1}^m (v_i^K + v_i^V)}. \quad (19)$$

The overall objective for the sustainable development of the regional ecosystem combines multiple subobjectives,

which are impossible to achieve all at once. To realize the overall objective, the traditional subobjective setting approach needs to adjust the subobjectives repeatedly. Based on multiple sets of environmental big data on the optimization system in different periods, this paper puts forward a novel weighted subobjective setting strategy.

According to the weights of subobjectives for the sustainable development of the regional ecosystem as well as the sustainable performance of each subobjective reflected by the multiple sets of environmental big data, the weighted subobjectives can be configured under a constraint as follows:

$$\begin{aligned} \max SG_i^* &= q_i \times \left[ \sum_{j \in T} (\psi_j \times \bar{\gamma}_{ij}^V) \right], \\ \text{s.t. } \psi_j(1 - \psi_j) &= 0, T \in D, \end{aligned} \quad (20)$$

where  $\bar{\gamma}_{ij}^V$  be the performance value of subobjective  $i$  reflected by the normalized environmental sample data  $j$  (If  $\psi_j = 1$  or 0 in  $\psi_j(1 - \psi_j) = 0$ , then the  $j$ -th set of environmental sample data can or cannot serve as a reference);  $D$  is the constraint;  $T \in D$  is the requirement that any combination of optimization measures must satisfy  $D$ .

Through the above steps, all weighted subobjectives can be obtained as  $SG^* = [SG_1^*, SG_2^*, \dots, SG_m^*]$ . Each weighted subobjective corresponds to a specific model. Each combination  $T_i^*$  of optimization measures leads to the optimal solution to each specific model.

This paper likens the set of optimal solutions on all weighted subobjectives as the ideal solution to the objective of sustainable development, which can be vectorized as

$$\vec{E}(JL) = [SG_1^*, SG_2^*, \dots, SG_m^*]. \quad (21)$$

The ideal solution above can be transformed to an actual solution under the condition  $T_1^* = T_2^* = \dots = T_m^*$ . For any

feasible solution, the corresponding combination of optimization measures  $T_b$  can be vectorized as

$$\begin{aligned}\vec{E}(T_b) &= [\vec{E}(T_b^K), \vec{E}(T_b^V)], \\ &= \{[SG_1(T_b^K), SG_1(T_b^V)], [SG_2(T_b^K), SG_2(T_b^V)], \dots, [SG_m(T_b^K), SG_m(T_b^V)]\}.\end{aligned}\quad (22)$$

Considering the length and direction of vector functions, the degree of absolute improvement of sustainable development in the ideal solution IS can be given by

$$N(\text{IS}) = \|\vec{E}(\text{IS})\| = \sqrt{\sum_{i=1}^m (SG_i^*)^2}. \quad (23)$$

The degree of absolute improvement of sustainable development in the feasible solution  $b$  can be given by

$$\begin{aligned}N(T_b) &= [N(T_b^K), N(T_b^V)] = [\|\vec{E}(T_b^K)\|, \|\vec{E}(T_b^V)\|], \\ &= \left[ \sqrt{\sum_{i=1}^m (SG_i(T_b^K))^2}, \sqrt{\sum_{i=1}^m (SG_i(T_b^V))^2} \right].\end{aligned}\quad (24)$$

The equilibrium of relative development of multiple weighted subobjectives in the ideal solution IS can be given by

$$\cos(F(\text{IS})) = \left( \frac{\vec{E}(\text{IS}) \cdot \vec{E}(\text{IS})}{\|\vec{E}(\text{IS})\| \cdot \|\vec{E}(\text{IS})\|} \right) = 1. \quad (25)$$

The equilibrium of relative development of multiple weighted subobjectives in the feasible solution  $b$  can be given by

$$\begin{aligned}\cos(F(T_b)) &= [\cos(F(T_b^K)), \cos(F(T_b^V))] = \left[ \min \left\{ \left( \frac{\vec{E}(\text{IS}) \cdot \vec{E}(T_b^K)}{\|\vec{E}(\text{IS})\| \cdot \|\vec{E}(T_b^K)\|} \right), \left( \frac{\vec{E}(\text{IS}) \cdot \vec{E}(T_b^V)}{\|\vec{E}(\text{IS})\| \cdot \|\vec{E}(T_b^V)\|} \right) \right\}, \right. \\ &\quad \left. \max \left\{ \left( \frac{\vec{E}(\text{IS}) \cdot \vec{E}(T_b^K)}{\|\vec{E}(\text{IS})\| \cdot \|\vec{E}(T_b^K)\|} \right), \left( \frac{\vec{E}(\text{IS}) \cdot \vec{E}(T_b^V)}{\|\vec{E}(\text{IS})\| \cdot \|\vec{E}(T_b^V)\|} \right) \right\} \right] \\ &= \left[ \min \left\{ \left( \frac{\sum_{i=1}^m [SG_i^* \times SG_i(T_b^K)]}{\sqrt{\sum_{i=1}^m (SG_i^*)^2} \times \sqrt{\sum_{i=1}^m [SG_i(T_b^K)]^2}} \right), \left( \frac{\sum_{i=1}^m [SG_i^* \times SG_i(T_b^V)]}{\sqrt{\sum_{i=1}^m (SG_i^*)^2} \times \sqrt{\sum_{i=1}^m [SG_i(T_b^V)]^2}} \right) \right\}, \right. \\ &\quad \left. \max \left\{ \left( \frac{\sum_{i=1}^m [SG_i^* \times SG_i(T_b^K)]}{\sqrt{\sum_{i=1}^m (SG_i^*)^2} \times \sqrt{\sum_{i=1}^m [SG_i(T_b^K)]^2}} \right), \left( \frac{\sum_{i=1}^m [SG_i^* \times SG_i(T_b^V)]}{\sqrt{\sum_{i=1}^m (SG_i^*)^2} \times \sqrt{\sum_{i=1}^m [SG_i(T_b^V)]^2}} \right) \right\} \right].\end{aligned}\quad (26)$$

By the principle of vector projection, the equilibrium of relative development of weighted subobjectives can be unified with the absolute improvement of sustainable development. In ideal conditions, the optimal degree of improvement of regional sustainable development can be given by

$$SP(\text{IS}) = [N(\text{IS}) \times \cos(F(\text{IS}))] = \sqrt{\sum_{i=1}^m (SG_i^*)^2}. \quad (27)$$

The improvement effect of the feasible solution  $T_b$  on sustainable development can be quantified by

$$SP(T_b) = [SP(T_b^K), SP(T_b^V)] = [N(T_b^K) \times \cos(F(T_b^K)), N(T_b^V) \times \cos(F(T_b^V))]. \quad (28)$$

The intelligent decision making of the objective of sustainable development can be derived from formulas (27) and (28). To display the decision-making process more intuitively, an objective achievement rate was introduced to transform the two formulas:

$$\text{OBJ}(T_l) = [\text{OBJ}(T_l^K), \text{OBJ}(T_l^V)] = \left[ \frac{SP(T_l^K)}{SP(\text{IS})}, \frac{SP(T_l^V)}{SP(\text{IS})} \right] \times 100\%. \quad (29)$$

The above formula shows that the objective achievement rate of IS is 1 and that of  $T_b$  falls in between (0, 1) (the closer the value is to 1, the better the optimization effect). Through the above analysis, the overall optimization objective of sustainable development for the regional ecosystem can be expressed as

$$\begin{aligned} & \text{Max OBJ}(T_l), \\ & \text{s.t. } T_l \in D. \end{aligned} \quad (30)$$

The above formula also outputs an interval number. Under uncertainty conditions, the model can be solved by probability comparison:

$$S_{(a>t)} = \mathbf{max} \left\{ 1 - \mathbf{max} \left( \frac{W_t^V - W_a^K}{W_t^V - W_a^K + W_t^V - W_a^K}, 0 \right), 0 \right\}. \quad (31)$$

Any interval numbers  $W_t = [W_t^K, W_t^V]$  and  $W_a = [W_a^K, W_a^V]$  can be compared in terms of probability. If  $W_a > W_b$ , i.e.,  $S_{(a>t)}, S_{(a>t)} > 0.5$  means  $W_a$  is probably better than  $W_b$ ; the inverse is also true. Similarly, if  $W_t^V < W_a^V$  and  $W_t^K < W_a^K$ ,  $W_a$  must be better than  $W_t$ .

#### 4. Experiments and Result Analysis

Figure 5 presents the evolution trends of the overall ecological efficiency of the study area in 2005–2020. The REEs exhibited obvious features of phased development, dropping from 0.6024 in 2006 to the lowest point of 0.5785 in 2013. From 2014 to 2020, the REEs rose significantly and peaked at 0.6549. The reasons for the phased development are as follows:

In the early phase of the sample period, the regional economy developed extensively in the traditional model, and the society had poor awareness of ecoenvironmental protection. That is why the REEs in the study area continued to decline. In 2013–2014, the cities in the study area quickly adjusted their development pattern and publicized ecological civilization, which contribute to the continuous growth of REEs.

There were some differences between regions in ecological efficiency. Region A, a transportation hub in a strategic location, remained the leader of ecological efficiency throughout the 15-year-long sample period. The ecological efficiencies of Regions A and B oscillated similarly as the overall ecological trend of the study area. After 2015, the ecological efficiency of Region B increased markedly each year, with increasingly strong momentum. Region C

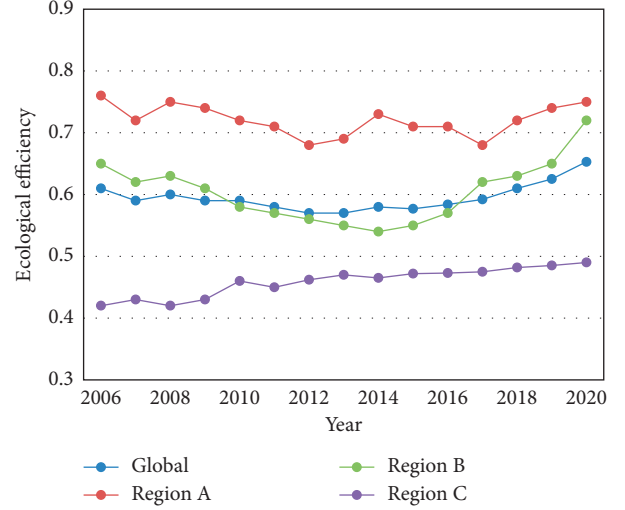


FIGURE 5: REE trends in 2005–2020.

had the lowest ecological efficiency, which grew stably over the 15 years.

Figure 6 compares the REE evolution trends of each pair of the three regions. Through the sample period, the largest regional difference in ecological efficiency existed between Regions A and B, whose mean Gini coefficient was 0.274. The second-largest regional difference was observed between Regions A and C, whose mean Gini coefficient was 0.261. The smallest regional difference was observed between Regions B and C, whose mean Gini coefficient was 0.253.

In general, the ecological efficiency gap between Regions A and C continued to narrow at an annual change rate of  $-2.14\%$ . The gap between Regions B and C widened after 2010, despite some fluctuations, with an annual change rate of  $0.87\%$ . Specifically, the gap between Regions A and C declined with fluctuations from 0.347 in 2006 to 0.206 in 2020, reaching the valley of the sample period. The gaps between Regions A and B, and between Regions B and C both decreased first and then increased. The gap between Regions A and B minimized at 0.223 in 2016, while that between Regions B and C minimized at 0.204 in 2010.

To sum up, the relevant government departments should strengthen the coordination between Regions B and C in green development and reduce the spatial difference between regions in ecological efficiency without ignoring the regional difference between A and B.

Through Gini coefficient decomposition, this paper looks for the sources of REE spatial differences, that is, analyzes the contributions of intraregion difference, inter-regional difference, and transvariation intensity to overall regional spatial differences. Table 2 presents the calculated results on intra and inter-regional contributions.

Through the sample period, the inter-regional difference of A and B contributed the greatest to the spatial gap of REE, followed in turn by that of B and C, and that of A and C; the intraregional difference of C contributed the greatest to the spatial gap of REE, followed in turn by that of A and that of B. It can be observed that the spatial imbalance of REEs can be effectively solved by reducing inter-regional differences.

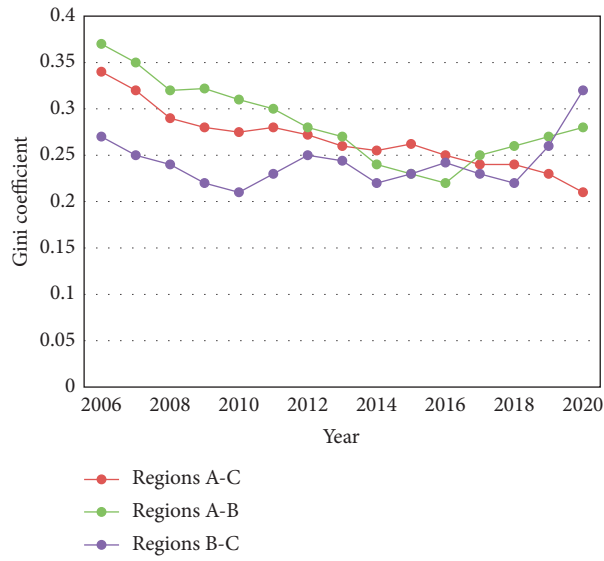


FIGURE 6: Trends of regional REE differences.

TABLE 2: Source of differences and contributions.

Year	Intraregional contribution			Inter-regional contribution		
	Region A	Region B	Region C	Region A	Region B	Region C
2005	12.36	11.95	6.35	7.86	17.38	11.24
2006	13.52	11.36	5.41	10.25	16.23	12.08
2007	14.75	10.27	5.23	11.75	19.75	10.61
2008	15.23	8.62	5.75	13.28	21.61	10.25
2009	14.76	10.15	5.76	10.94	20.85	14.31
2010	14.76	10.36	5.23	8.51	17.23	14.52
2011	15.23	10.62	4.81	8.72	16.05	11.23
2012	15.08	13.54	5.08	7.63	17.28	13.54
2013	13.27	15.76	3.64	3.55	15.36	17.21
2014	10.75	13.28	4.76	3.42	17.54	20.36
2015	11.25	11.39	4.52	0.46	18.95	20.85
2016	13.62	14.25	4.08	5.23	21.72	28.30
2017	11.22	13.21	2.59	7.41	19.35	27.26
2018	9.35	15.02	2.79	2.88	20.08	25.42
2019	11.76	11.53	1.26	6.58	22.72	27.36
2020	10.05	9.74	2.65	6.75	23.45	31.65
Mean	12.94	11.94	4.37	7.21	19.09	14.76

Regions A and B should be the focus of regulation because their inter-regional difference is the major contributor to the spatial gap of REE.

Tables 3–5 report the transition probability matrices of the Markov chain for Regions A–C, respectively. Note that period  $p_1$  has a one-year lag, and period  $p_2$  has a two-year lag. The long-term transition trend of REE can be derived from these matrices. In each matrix, the elements to the left of the diagonal were smaller than those to the right, indicating that the overall ecological efficiency of each region moves clearly from the low level to the high level. Meanwhile, the REE had a greater stable probability than transition probability. This means that the state transition of ecological efficiency is accompanied by apparent polarization. As period  $p_1$  changed to period  $p_2$ , i.e., the lag increased from 1 year to 2 years, the overall ecological efficiency of the study area

basically remained the same, with a very slight decline. In addition, there were many nonzero elements on both sides of the diagonal in each matrix, suggesting the difficulty for the study area to realize the leapfrog transition of ecological efficiency.

This paper employs the multistage method to measure the contributions of intraregion difference, inter-regional difference, and transvariation intensity in 2005–2020. The results are presented in Table 6 and Figure 7. The measured results of our method have a smaller variation than those of the traditional DEA. The following can be inferred from the data in Table 6. In 2005–2020, the mean contribution was 0.813, 0.990, and 0.821 for intraregional difference, inter-regional difference, and transvariation intensity, respectively. Overall, the study area had relatively high ecological efficiency since 2005. The regional economic development



TABLE 3: Transition probability matrix of the Markov chain for Region A.

Period Type	$p_1$				$p_2$			
	1	2	3	4	1	2	3	4
1	0.7235	0.1122	0	0	0.6758	0.1523	0	0
2	0.2753	0.6235	0.0851	0	0.1532	0.5123	0.1231	0
3	0	0.2753	0.6723	0.1123	0	0.3154	0.5876	0.1531
4	0	0	0.2235	0.8517	0	0	0.2753	0.8675

TABLE 4: Transition probability matrix of the Markov chain for region B.

Period Type	$p_1$				$p_2$			
	1	2	3	4	1	2	3	4
1	0.5321	0.0223	0	0	0.3725	0.0275	0	0
2	0.4233	0.4555	0.0853	0.0153	0.5742	0.0375	0.0723	0.0181
3	0.0257	0.4125	0.0545	0.1535	0.0436	0.3122	0.4368	0.2376
4	0.0257	0.1333	0.3721	0.8755	0.0214	0.2763	0.4214	0.7532

TABLE 5: Transition probability matrix of the Markov chain for region C.

Period Type	$p_1$				$p_2$			
	1	2	3	4	1	2	3	4
1	0.6355	0.1321	0	0	0.5728	0.0972	0	0
2	0.3721	0.5765	0.1123	0	0.3721	0.5321	0.1423	0.0354
3	0.0253	0.2675	0.5721	0.2222	0.0434	0.3122	0.5472	0.2235
4	0	0.0223	0.3555	0.7655	0	0.0251	0.3222	0.7652

TABLE 6: Overall ecological efficiencies of the study area in 2005–2020.

Year	Intraregional difference	Inter-region difference	Transvariation intensity
2005	0.932	0.915	1
2006	1	1	1
2007	0.975	0.962	1
2008	0.982	0.988	0.962
2009	0.976	0.975	1
2010	0.752	0.752	1
2011	0.857	0.813	1
2012	0.753	0.756	1
2013	0.769	0.769	0.988
2014	0.982	0.982	1
2015	0.974	0.974	1
2016	0.993	0.993	1
2017	0.871	0.854	1
2018	0.963	0.921	0.994
2019	0.842	0.867	1
2020	0.873	0.893	1

achieves a high cost effectiveness in terms of energy-saving technology and green environmental investment. In 2010 and 2013, the contribution of inter-regional difference was merely 0.758 and 0.761, respectively. In 2009–2011 and 2013–2016, the inter-region contributions

were below the mean of the 15 years. Through the 15-year-long sample period, the transvariation intensity made relatively high contributions, reflecting the good effect of technical and management measures adopted by the regional ecosystem.

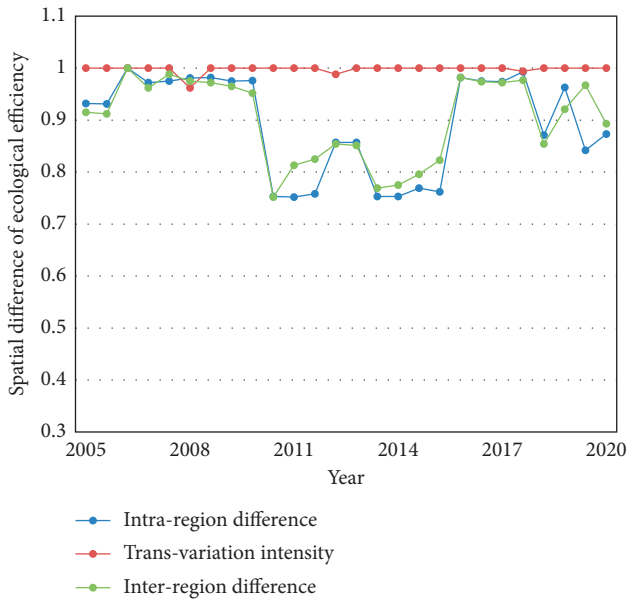


FIGURE 7: REE decomposition and changes in 2005–2020.

## 5. Conclusions

Based on the analysis of environmental big data, this paper investigates the evaluation of REE and intelligent decision support for sustainable development. Specifically, the spatiotemporal evolution of REE was analyzed based on environmental big data, and the spatial layout of REE and the sources of the spatial differences were both clarified. In addition, the authors constructed a multiobjective optimal decision-making model for the sustainable development of a regional ecosystem and presented the solving method for the model. After that, experiments were carried out to analyze the REE changes and inter-regional REE differences in 2005–2020. The experimental results demonstrated the feasibility of our analytical method. Furthermore, the authors discussed the contributions of intraregional difference, inter-regional difference, and transvariation intensity to the overall spatial difference of regional ecological efficiency, constructed the transition probability matrix of the Markov chain for the REE in each region, and provided the measured results on regional REEs.

The future research will try to develop a decision model capable of considering and solving different interest preferences and demands of different decision makers and realize multirole, multiattribute decision making as well as multirole, multiobjective decision making.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported by the Humanities and Social Science Foundation of Education Department of Hubei Province (Grant no. 194101), the Young Scientists Fund of Humanities and Social Sciences of the Ministry of Education of China (Grant no. 19YJC630236), and the Philosophy and Social Science Research Project of the Hubei Provincial Department of Education (Grant no. 19Q093).

## References

- [1] D. S. Schodermayr, K. P. Lafayette, J. Bezerra, D. F. da Paz, R. R. Sousa, and R. Albuquerque, "Analysis of the environmental impacts arising from irregular civil construction waste disposal in the city of Jaboatao dos Guararapes/PE," *European Journal of Government and Economics*, vol. 20, pp. 12791–12800, 2015.
- [2] S. C. Swarnokar, M. Ashik-Ur-Rahman, and S. I. Mou, "Conflict of resource use among different livelihood group in coastal villages of south-western Bengal Delta, Bangladesh," *International Journal of Sustainable Development and Planning*, vol. 15, no. 7, pp. 1089–1099, 2020.
- [3] G. M. E. Kamh, "Environmental impact on construction limestone at humid regions with an emphasis on salt weathering, Al-hambra islamic archaeological site, Granada City, Spain: case study," *Environmental Geology*, vol. 52, no. 8, pp. 1539–1547, 2007.
- [4] Y. Tang and X. Deng, "Economic and environmental impacts of geothermal resource development in Hunan, China," *International Journal of Heat and Technology*, vol. 39, no. 2, pp. 581–586, 2021.
- [5] J. K. Ssegawa, P. Mselle, M. Matlhare, and C. Ditshane, "Construction activities and environmental issues in and around Gaborone city," *International Journal of Environmental Technology and Management*, vol. 2, no. 1-3, pp. 127–141, 2002.
- [6] H. Wang and S. Cheng, "Spatiotemporal variation in land use of northeast China tiger and leopard national park," *International Journal of Design & Nature and Ecodynamics*, vol. 15, no. 6, pp. 835–842, 2020.
- [7] L. Li and S. Zhang, "Techno-economic and environmental assessment of multiple distributed energy systems coordination under centralized and decentralized framework," *Sustainable Cities and Society*, vol. 72, Article ID 103076, 2021.
- [8] H. Han, L. Guo, J. Zhang, K. Zhang, and N. Cui, "Spatio-temporal analysis of the coordination of economic development, resource utilization, and environmental quality in the Beijing-Tianjin-Hebei urban agglomeration," *Ecological Indicators*, vol. 127, Article ID 107724, 2021.
- [9] P. Liu, Z. Fang, C. Lv, and A. Ruan, "China's agricultural water-use efficiency and its influencing factors under the constraint of pollution emission," *International Journal of Design & Nature and Ecodynamics*, vol. 15, no. 4, pp. 579–585, 2020.
- [10] M. Zare Oskouei, B. Mohammadi-Ivatloo, M. Abapour, M. Shafiee, and A. Anvari-Moghaddam, "Techno-economic and environmental assessment of the coordinated operation of regional grid-connected energy hubs considering high penetration of wind power," *Journal of Cleaner Production*, vol. 280, Article ID 124275, 2021.
- [11] M. E. Nazari and M. M. Ardehali, "Optimal coordination of renewable wind and pumped storage with thermal power generation for maximizing economic profit with

- considerations for environmental emission based on newly developed heuristic optimization algorithm,” *Journal of Renewable and Sustainable Energy*, vol. 8, no. 6, Article ID 65905, 2016.
- [12] D. Peng, Y. Mu, and Y. Zhu, “Evaluating the level of coordinated development of fisheries economic growth and environmental quality in selected Chinese regions,” *Environmental Impact Assessment Review*, vol. 89, Article ID 106605, 2021.
- [13] G. Yao, M. Wang, Y. Yang, and J. Li, “Development and analysis of prefabricated concrete buildings in Chengdu, China,” *International Journal of Sustainable Development and Planning*, vol. 15, no. 3, pp. 403–411, 2020.
- [14] M. D. Ibrahim, A. A. Alola, and D. Cunha Ferreira, “A two-stage data envelopment analysis of efficiency of social-ecological systems: inference from the sub-Saharan African countries,” *Ecological Indicators*, vol. 123, Article ID 107381, 2021.
- [15] L. O. Soares, D. R. de Moraes, and R. A. M. Boloy, “Energy-ecological efficiency and TTW emissions of the DFSIE fuelled with biofuels,” *BioEnergy Research*, vol. 14, no. 2, pp. 623–633, 2021.
- [16] T. Mahmood, S. Shireen, and M. Mumtaz, “Testing the role of financial development and urbanization in the conventional EKC: evidence from China and India,” *International Journal of Sustainable Development and Planning*, vol. 16, no. 3, pp. 445–455, 2021.
- [17] B. Han, Z. Ouyang, and W. Wang, “The relationship between regional industrial organizing levels and ecological economic efficiency,” *Journal of Cleaner Production*, vol. 171, pp. 857–866, 2018.
- [18] J. Nowakowski, K. Brzozowski, and T. Knefel, “Controlling efficiency of engine’s adjustment parameters in relation to ecological criteria,” in *Proceedings of the Transport Means-Proceedings of the International Conference*, pp. 606–609, Trakai, Lithuania, October 2018.
- [19] D. A. Rojas-Gamboa, J. I. Rodriguez, J. Gonzalez-Ayala, and F. Angulo-Brown, “Ecological efficiency of finite-time thermodynamics: a molecular dynamics study,” *Physical Review*, vol. 98, no. 2, Article ID 22130, 2018.
- [20] J. Ahmed, A. Saeed, and A. Saeed, “Urban sustainability between administrative decision-making and the environment source,” *International Journal of Sustainable Development and Planning*, vol. 15, no. 5, pp. 713–720, 2020.
- [21] L. E. Shvartsburg, E. V. Butrimova, and O. V. Yagolnitsner, “Energy efficiency and ecological safety of shaping technological processes,” *Procedia Engineering*, vol. 206, pp. 1009–1014, 2017.
- [22] L. Liu and Q. Sun, “Empirical research on ecological efficiency of coal resource-dependent cities in China,” *Journal of Environmental Engineering*, vol. 145, no. 9, Article ID 4019047, 2019.
- [23] S. V. Ratner, “Dynamic problems of estimation of ecological-economic efficiency of regions based on basic models of data envelopment analysis,” *Automation and Remote Control*, vol. 80, no. 12, pp. 2245–2258, 2019.
- [24] P. Giordano, P. Caputo, and A. Vancheri, “Fuzzy evaluation of heterogeneous quantities: measuring urban ecological efficiency,” *Ecological Modelling*, vol. 288, pp. 112–126, 2014.
- [25] V.-N. Hoang and M. Alauddin, “Input-orientated data envelopment analysis framework for measuring and decomposing economic, environmental and ecological efficiency: an application to OECD agriculture,” *Environmental and Resource Economics*, vol. 51, no. 3, pp. 431–452, 2012.
- [26] J. Wang and X. P. Li, “The sustainable space development model based on smart urban,” *Chemical Engineering Transactions*, vol. 51, pp. 853–858, 2016.
- [27] M. S. Ma, M. Zhao, J. Y. Chen, and J. L. Li, “Research on sustainable development capacity for the urban ecosystem—a case study,” *Chemical Engineering Transactions*, vol. 51, pp. 817–822, 2016.

## Research Article

# A Big Data-Driven Approach to Analyze the Influencing Factors of Enterprise's Technological Innovation

**Qianqian Zhang** 

*School of Information, Beijing Wuzi University, Beijing 101149, China*

Correspondence should be addressed to Qianqian Zhang; [zhangqianqian@bwu.edu.cn](mailto:zhangqianqian@bwu.edu.cn)

Received 20 November 2021; Accepted 4 January 2022; Published 25 January 2022

Academic Editor: Daqing Gong

Copyright © 2022 Qianqian Zhang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A data-driven intelligent analysis method is proposed in this paper to explore and identify the enterprise's technological innovation influencing factors. Questionnaire surveys or expert interviews are usually adopted by the traditional evaluation methods for indicators of technological innovation selection. However, it inevitably involves human factors and experts' subjective judgments, which may affect the result of enterprises evaluation. The research presents an improved text clustering method based on a semantic concept model to explore and analyze the key influencing factors of enterprise's technological innovation. The study collects textual data from 400 enterprises in Beijing and smart analyzes the critical influencing factors of enterprise's technological innovation by using the proposed method. The influencing factors can be divided into seven categories. In addition, compared with the traditional K-means clustering method, the proposed method has a good effect. We proposed a methodology to conduct an intelligent analysis for enterprise's technological innovation under the data-driven. It can provide more objective and auxiliary suggestions for the evaluation of the enterprise's technology innovation.

## 1. Introduction

Technological innovation is the foundation of the survival and development of enterprises and the driving force for the country's economic and social development. It is essential for an enterprise to gain a competitive advantage by correctly analyzing and evaluating the technological innovation capability. Since the middle of the 18th century, the world has experienced three industrial revolutions. Social development is entering the era of data revolution with the development and application of new-generation information technologies such as cloud computing, big data, mobile Internet, artificial intelligence, and the Internet of Things. The production and circulation of massive data gave birth to "big data" and set off the fourth industrial revolution.

The digital economy era has given birth to new production factors represented by big data. Data-driven continuous growth and innovative development are the main lines of enterprise digital transformation. Compared with the past process-driven, data-driven allows companies to use massive and multidimensional data to establish a more

comprehensive evaluation system, create direct business innovation growth, and continuously improve operational efficiency. It is an essential means to maintain sustainable development in market competition. With the continuous development of enterprises, various forms of technical documents and text information continue to spew. According to statistics, 80% of enterprise data exist in unstructured conditions, such as Web pages, technical papers, e-mails, etc. Especially in enterprise's technological innovation, related technological innovation activity reports, meeting minutes, annual corporate reports, patent technology files, project application reports, and textual information are increasing with each passing day. Most of the time, enterprises often need to deal with disordered unstructured textual data except structured data. Ignoring the text information generated by corporate technological innovation activities will inevitably affect the result of technical innovation management. Enterprises need to deal with these collected files and explore the value and knowledge behind these massive amounts of data. Therefore, the rise of big data and the development of intelligent text analysis technology

have used many unstructured and fragmented textual data that enterprises initially neglected. However, it is a significant challenge for the era of big data to mine knowledge from massive unstructured data and provide auxiliary decision-making for enterprise technological innovation.

Text mining generally refers to the process of extracting valuable, nontrivial patterns or knowledge from a large amount of unstructured or semi-structured text files. Therefore, text mining provides an effective measure for textual data collation, analysis and, mining. It meets the massive demand for processing and analyzing an enormous amount of unstructured and semi-structured textual data. To a certain extent, it solves the labour and material cost problems of manual text processing. Text mining, as a representative of the intelligent measure method, has been widely applied in various fields. For example, the evaluation of business intelligence and enterprise technology analysis [1–3], the enterprise technology opportunity identification [4, 5], correlation analysis of enterprise technology cooperation behavior [6–8], the analysis of enterprises technology maturity [9–11], and prediction of enterprise technology development trend, etc. [12–14].

The intelligent text mining algorithm can quickly and high-quality organize amounts of information into a few meaningful clusters and obtain the hidden potential knowledge or patterns. The rapid growth of textual data, especially in enterprise's technological innovation, has become diverse, high-dimensional, complex data loaded with semantic information. Therefore, it is feasible to explore the essential influencing factors of enterprise's technological innovation based on intelligent analysis algorithms and realize the semantic organization for technological innovation evaluation. This paper collects technical data from 400 enterprises in Beijing. It combines the proposed intelligent text clustering algorithm to realize knowledge mining and acquire enterprise's technological innovation at the semantic level.

The remainder of this paper is organized as follows. The second section presents the related works. The third section provides the implementation process of the proposed methodology and makes a performance comparison validation with the traditional K-means clustering method. While section four presents the result and the section five give a discussion of this research. In the final section, we conclude this paper's work.

## 2. Related Works

The previous studies of enterprise's technological innovation usually focus on theory research and evaluation methods. In the early stage, the evaluation system of enterprise's technological innovation was mostly constructed by questionnaire or expert interview. The influencing factors of an enterprise's technological innovation capability were obtained by the questionnaire design or expert experience. However, these methods are often affected by the limitation of the sample size and the subjective factors of expert opinions. It is difficult to objectively and comprehensively reflect the status of enterprise's technological innovation

capabilities. Many scholars make a comprehensive evaluation of enterprise technological innovation. They usually use various evaluation methods to express the overall characteristics of enterprise technological innovation. These evaluation methods are mainly divided into the following aspects.

Some scholars adopted the fuzzy evaluation method to evaluate the enterprise technological innovation ability. The main feature of this method is first to design a set of evaluation index systems, determine the weight of each index, establish a fuzzy comment set, and then use the fuzzy evaluation method to judge the innovation ability of the enterprise. For example, Du et al. [15] established a risk evaluation model for technological innovation based on fuzzy evaluation. Suder and Kahraman [16] proposed a Fuzzy TOPSIS method to evaluate technological innovation investments using eight different criteria. Feng and Ma [17] identified the influencing factors of service innovation in manufacturing enterprises by using the fuzzy DEMATEL method. The shortcomings of this type of method are that human factors play a prominent role, and the data collection and processing are human-oriented, which lack objectivity and needs a lot of labour.

Some scholars used the Analytic Hierarchy Process (AHP) and its variants to evaluate the enterprise technological innovation capability. Mu et al. [18] established an index system for the technological innovation capabilities of small and medium-sized enterprises through the AHP method. Pan et al. [19] combined the AHP and osculating value process (OVP) to evaluate the green innovation ability of manufacturing enterprises. In addition, some scholars have used the improved AHP and fuzzy evaluation method to establish a model for assessing the technological innovation capabilities of enterprises [20]. The weight of each indicator in the AHP method depends on the subjective judgment of experts, and it is inevitable to have a certain degree of subjectivity.

Some researchers evaluated the enterprise technological innovation capability with the Data Envelopment Analysis (DEA) method. Wang et al. [21] constructed a high-tech industrial evaluation framework of technical innovation efficiency based on two-stage network data envelopment analysis (DEA). Ma et al. [22] used the DEA method to evaluate and analyze the innovation capability of 233 listed companies in 5 major industrial sectors defined by the China Securities Regulatory Commission (CSRC). Li et al. [23] measured the technical efficiency, scale efficiency, and pure technical efficiency of innovation in China's semiconductor industry using a three-stage DEA model. Although the DEA method can realize the multiple inputs and output, it is only an efficiency evaluation method and cannot indicate the actual technical level of the research object.

Some scholars proposed to use intelligent decision-making methods to evaluate the technological innovation capabilities of enterprises. The intelligent decision-making method applied the artificial intelligence-related theoretical methods and fusing traditional decision-making mathematical models for intelligent reasoning and solving, such as genetic algorithm, ant colony algorithm, rough set, and so



on. At present, there is little literature on the application of intelligent decision-making methods to evaluate enterprise technological innovation capabilities. Shang [24] proposed the evaluation model of strategic management capability based on the Back-Propagation (BP) neural network algorithm. Zhen and Yao [25] analyzed the lean production and technological innovation in the manufacturing industry based on Support Vector Machine (SVM) algorithms and data mining technology.

Most of the related works about enterprise's technological innovation evaluation methods rely on the experts' subjective experience. Hence, the evaluation results are varied due to the different experts' opinions. Less-comprehensive impact indicators selected by experts may not reflect the actual status of the enterprise's technological innovation capability. The lack of objectivity in the evaluation method will not be conducive for identifying and cultivating enterprise's technological innovation capabilities. At present, it can objectively reflect the actual level of enterprise's technological innovation capabilities by applying big data-driven intelligent analysis methods. In addition, it can use massive and multidimensional data to establish a more comprehensive evaluation system.

### 3. Methodology

This paper proposed an improved semantic clustering algorithm combined with the concept of semantic similarity and relatedness based on domain ontology. After text preprocessing, a corpus containing the keywords was built and the keywords were mapped to the concepts in the domain ontology of the enterprise's technological innovation. The calculation of semantic similarity and relatedness between concepts was one of the critical steps. It needs to use the semantic similarity and relatedness method proposed in this paper to establish the compound similarity matrix. Finally, the keyword set in the corpus is clustered according to the improved semantic text clustering algorithm proposed in this paper. Figure 1 shows the framework of semantic-based text clustering in the field of enterprise technology innovation. According to the framework, there are mainly three steps in the improved methodology and the details are as follows, and the main parameters used in the following equations are shown in Table 1.

**3.1. Text Concept Mapping.** After collecting the data from the enterprise's technological innovation, preprocessing the data, including Chinese word segmentation, custom dictionary, POS selection, and stop words removing, etc., obtained the text keyword sets. The keyword set is mapped to the ontology of the enterprise technology innovation domain to obtain the concept set. Two situations need to be considered.

- (1) When the dataset's keywords can directly match with the concepts in the domain ontology, the keywords  $T = \{t_1, t_2, \dots, t_n\}$  are directly mapped with the concepts  $C = \{c_1, c_2, \dots, c_m\}$ .

- (2) When the keywords in the dataset cannot directly match the concepts in the domain ontology and appear frequently, the keywords should be reserved as unregistered words. Calculate the occurrence frequency TF of the keyword. If  $TF > \mu$ , keep the keyword in the unregistered word sets  $W = \{w_1, w_2, \dots, w_l\}$ , otherwise delete the keyword.

**3.2. Semantic Similarity and Relatedness Calculation.** Before text clustering, the semantic similarity and relatedness calculation method need to be used to construct the semantic matrix. We proposed a new semantic measurement method that combines the concepts of semantic similarity and relatedness. Firstly, calculate the semantic similarity of concepts based on the semantic distance in the established domain ontology of enterprise's technological innovation, as shown in Figure 2. Secondly, assign the weights to the path of two concepts connected in the domain ontology. The semantic distance between two concepts was obtained by traversing the sum of the weights of the connection paths instead of calculating the number of edges connecting the two concepts. The specific calculation is shown in equation (1). Then, calculate the semantic relatedness of concepts through the co-occurrence in the text. Finally, combine the semantic similarity and the relatedness to establish the compound similarity matrix  $M$ .

$$M = \begin{bmatrix} 1 & \dots & \text{sim\_rel}_{i1} & \dots & \text{sim\_rel}_{j1} & \dots & \text{sim\_rel}_{l1} \\ \vdots & \ddots & \dots & \vdots & \dots & \vdots & \vdots \\ \text{sim\_rel}_{i1} & \dots & 1 & \dots & \text{sim\_rel}_{ji} & \dots & \text{sim\_rel}_{li} \\ \vdots & \vdots & \dots & \ddots & \dots & \vdots & \vdots \\ \text{sim\_rel}_{j1} & \dots & \text{sim\_rel}_{ji} & \dots & 1 & \dots & \text{sim\_rel}_{lj} \\ \vdots & \vdots & \dots & \vdots & \dots & \ddots & \vdots \\ \text{sim\_rel}_{i1} & \dots & \text{sim\_rel}_{i1} & \dots & \text{sim\_rel}_{lj} & \dots & 1 \end{bmatrix}, \quad (1)$$

$$\begin{aligned} \text{Sim\_Rel}(c_i, c_j) &= \alpha \times \text{sim}(c_i, c_j) \\ &+ (1 - \alpha) \times \text{rel}(c_i, c_j) \\ &= \alpha \times \frac{1}{1 + \lambda \text{dist}(c_i, c_j)} + (1 - \alpha) \times \log 2 \frac{f^k(c_i, c_j)}{f^c(c_i) \times f^c(c_j)}, \end{aligned}$$

where  $\alpha$  denotes the adjusting parameter,  $\lambda$  denotes the factor of influence degree of semantic distance on semantic similarity,  $\text{dist}(c_i, c_j)$  denotes the semantic distance between  $c_i$  and  $c_j$  in the domain ontology,  $f^k$  denotes the co-occurrence frequency of concept  $c_i$  and concept  $c_j$  in the  $k$  words window of the entire corpus.  $f^c(c_i)$  and  $f^c(c_j)$  represent the frequency of concept  $c_i$  and concept  $c_j$  in the whole corpus.

**3.3. Improved K-Means Algorithm.** The traditional clustering algorithm based on vector space defines the text document set as  $D = \{d_1, d_2, d_3, \dots, d_m\}$  and each document can be



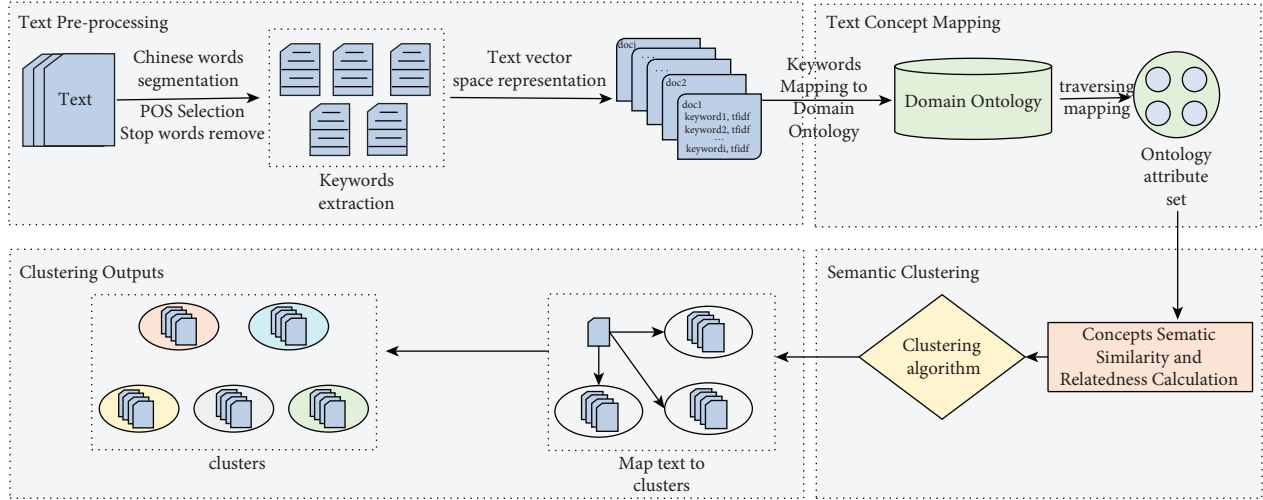


FIGURE 1: Implementation process of text clustering method based on semantic similarity and relatedness.

TABLE 1: The main parameters in equations' definitions.

Elements	Definition
$t_n$	Text keywords
$c_m$	Concepts in the domain ontology
$w_l$	Unregistered word
TF	The frequency of the keyword in dataset
$\mu$	The threshold value of keyword frequency
$M$	The semantic similarity and the relatedness matrix
$\text{dist}(c_i, c_j)$	The semantic distance between concepts $c_i$ and $c_j$ in ontology
$\lambda$	The influence factor of semantic distance on semantic similarity
$f^k(c_i, c_j)$	The number of times that concept $c_i$ and $c_j$ appear simultaneously in the $k$ words window at the entire corpus
$f^c(c_i)$	The frequency of the concept $c_i$ at the entire corpus
$\text{rel}(c_i, c_j)$	Relatedness between concept $c_i$ and $c_j$
$\text{Sim\_Rel}(c_i, c_j)$	Semantic similarity and relatedness between concept of $c_i$ and $c_j$
$\alpha$	The adjusting parameter
$d_i$	The text document
$w_{ij}$	The weight of the term $t_j$ in the document $d_i$
$t_j$	The feature vector of the document
$z_l$	The cluster centre
$\hat{t}_j$	The new document feature vector derived from $t_j$
$\bar{z}_l$	The cluster centre derived from $z_l$

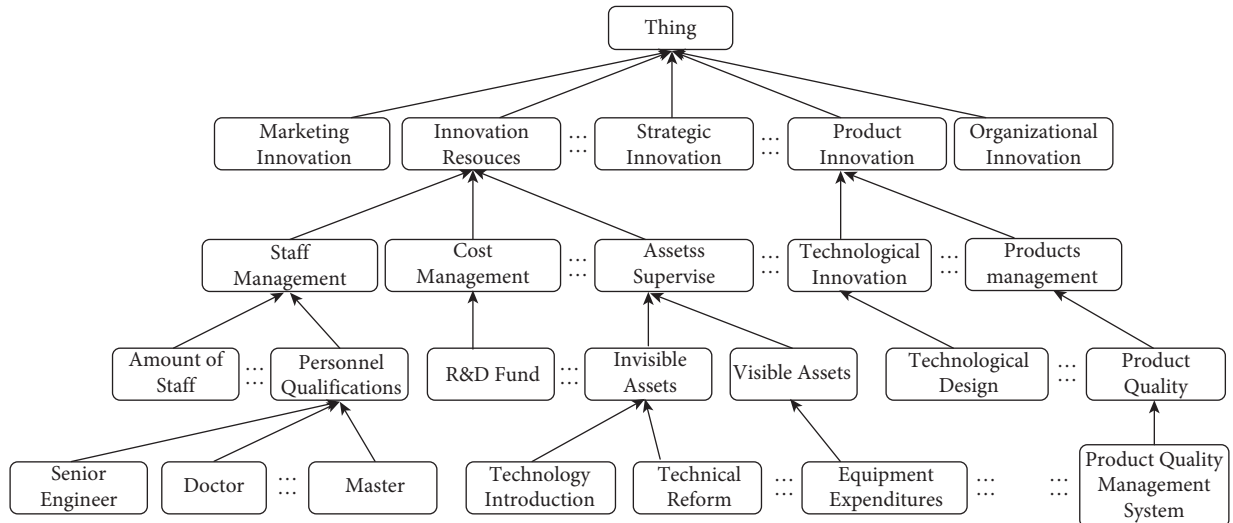


FIGURE 2: Structure of partial enterprise's technological innovation domain ontology.

represented as  $d_i = (w_{i1}, w_{i2}, w_{i3}, \dots, w_{in})$ . The term set extracted from the document set can be expressed as  $T = \{t_1, t_2, t_3, \dots, t_n\}$ .  $w_{ij}$  means the weight of the term  $t_j$  in the document  $d_i$ . The traditional clustering algorithms are usually determined by the number of occurrences of the term  $t_j$  appearing in the document  $d_i$ , that is, the term frequency, or the term frequency-inverse document frequency (TF-IDF), is used to assign weights. The traditional K-means algorithm selected  $K$  points randomly from the sample as the initial cluster centre candidate. Then, it usually calculated the distance from the sample points to the centre with the Euclidean distance formula (shown in equation (2)) and divided the points to the nearest centre. Finally, it iteratively calculated the centre of the cluster until the centre of each group does not change.

$$\begin{aligned} \text{dis}_{EU}(d_1, d_2) &= \sqrt{(d_1 - d_2)(d_1 - d_2)^T} \\ &= \sqrt{\sum_{j=1}^n (w_{1j} - w_{2j})^2}. \end{aligned} \quad (2)$$

This paper proposed an improved K-means algorithm, which improves the algorithm mainly by selecting initial cluster centres and semantic-based clustering. According to the semi-positive semantic similarity and relatedness of  $n \times n$  matrix  $M$  obtained by Step3,  $M = W^T W$  can be obtained by orthogonalization of the positive semi-positive matrix, where the column in  $W$  represents the document vector. According to the semantic similarity and relatedness matrix  $M$ , the Euclidean distance formula can be improved as the following equation.

$$\begin{aligned} \text{dis}_{EU \text{ improved}}(d_1, d_2) &= \sqrt{(d_1 - d_2)(M d_1 - d_2)^T} \\ &= \sqrt{(d_1 - d_2)W^T W (d_1 - d_2)^T} \\ &= \sqrt{((d_1 - d_2)W)((d_1 - d_2)W)^T} \\ &= \sqrt{(\hat{d}_1 - \hat{d}_2)(\hat{d}_1 - \hat{d}_2)^T}. \end{aligned} \quad (3)$$

where  $\hat{d}_1 = d_1 W$ ,  $\hat{d}_2 = d_2 W$ , the modified distance measurement considered the semantics between words based on the semantic similarity and relatedness matrix. Therefore, the distance between the document and the cluster can be measured by the definition of the following equation:

$$\begin{aligned} \text{dis}_{EU \text{ improved}}(t_j, z_l) &= \sqrt{(t_j W - z_l W)(t_j W - z_l W)^T} \\ &= \sqrt{(\hat{t}_j - \hat{z}_l)(\hat{t}_j - \hat{z}_l)^T}, \end{aligned} \quad (4)$$

where  $t_j$  represents the feature vector of the document,  $z_l$  defines the cluster centre,  $\hat{t}_j$  is the new document feature vector derived from  $t_j$ , and  $\hat{z}_l$  is the cluster centre derived from  $z_l$ .

Firstly, randomly select a data point  $z_l$  as the initial cluster centre from the input dataset. Secondly, for each point  $c_i$  in the dataset, the distance from  $c_i$  to  $z_l$  is calculated

by the Equation  $D(c_i) = \arg \min \|c_i - z_j\|_2^2, j = 1, 2, \dots, k$ . The point  $c_i$  with the maximum distance  $D(c_i)$  from  $z_l$  is selected as the new cluster centre  $z_2$ . Then, repeat the above steps until  $K$  initial cluster centres are found. Finally, the remaining data points in the sample set are allocated to the nearest or most similar clusters according to the principle of maximum similarity. The cluster centres in the  $K$  clusters are recalculated and iterated until the termination condition is met. The step of the improved K-means algorithm is described in Algorithm 1.

The improvement to the traditional K-means algorithm proposed in this paper based on semantic similarity and relevance mainly includes (1) using the improved method based on the maximum similarity to determine the initial cluster centre and reduce the random position of the cluster centre. (2) The improved Euclidean distance with semantic similarity and relatedness is used to measure the similarity between cluster centres and sample sets, instead of the traditional K-means algorithm, which ignores the semantic relationship between terms. (3) By adding convergence conditions, the original K-means algorithm solved the problem of unstable clustering results.

### 3.4. Validation

**3.4.1. Validation Methods.** The paper mainly used the SSE and SC methods to compare the clustering results of the traditional K-means method and the improved K-means method. The SSE method calculates the total error value between any data point and the cluster centroid, and the calculation method is shown in equation (5).  $\text{dis}$  represents the distance function,  $p$  is any data point in the cluster of  $c_i$ , and  $m_i$  is the cluster centroid. The lower value of SSE equates to better performance of clustering. Otherwise, a higher value represents a worse clustering effect. The Silhouette Coefficient method combines clustering cohesion and separation to evaluate the effect of clustering, and the value is between  $[-1, 1]$ . The higher value of the Silhouette Coefficient indicates the better clustering performance. The calculation method of Silhouette Coefficient was shown in equation (6).  $a(i)$  represents the average distance of the data point  $i$  to all other points in the cluster to which the data point  $i$  belongs.  $b(i)$  represents the minimum value of the average distance from the data point  $i$  to all points of each of the other groups.

$$\text{SSE} = \sum_{i=1}^k \sum_{p \in c_i} \text{dis}(p, m_i)^2, \quad (5)$$

$$S(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}. \quad (6)$$

**3.4.2. Validity Comparison of the Proposed Method.** To compare the clustering performance between the traditional K-means algorithm based on Bag-of-words and the improved K-means algorithm based on semantic similarity and relatedness, the number of  $K$  is selected from 3 to 10. The

**Input:** preprocessed dataset  $D = \{d_j | j = 1, 2, \dots, m\}$ ; the dataset contains  $N$  terms  $C = \{c_i | i = 1, 2, \dots, n\}$ , semantic similarity, and relatedness matrix  $M$ ; the number of cluster  $K$ ; iteration termination condition  $\epsilon$ ; the maximum number of iterations  $\text{MaxStep}$ ;

**Output:**  $K$  cluster result;

**BEGIN**

- (1) start = 0  
 $k = 0$ ; //initialization  
load dataset  $D$  and select an initial cluster centre  $z_1$  randomly from  $D$  saving to the initial cluster centre  $Z_j = \{z_j, j = 1, 2, \dots, k\}$ ;
- (2) Calculate the distance between each sample and the initial point  $z_l$ , find the point  $c_i$  with the largest distance from  $z_l$  according to the equation (4), take the sample point  $c_i$  as the second initial cluster centre  $z_2$ , and save it to the initial point set  $Z_j = \{z_j, j = 1, 2, \dots, k\}$ ;
- (3) repeat step 2 until the  $k$ th initial cluster centre is found;
- (4) according to the  $D(c_i) = \arg \min \|c_i - z_j\|_2^2$ , assign each sample  $c_i$  to the class of the nearest  $k$  initial cluster centres;
- (5) update the centre of each cluster through the mean value  $Z_i(O) = \sum_{i=1}^{n_i} c_i/n_i$ ,  $n_i$  represents the number of sample points in the group;
- (6) the measure function  $E = \sum_{l=1}^k \sum_{j=1}^m \sum_{i=1}^{n_i} W_{lj} \text{dis}_{EU \text{ improved}}(\tilde{z}_{li}, \hat{c}_{ji})$ ,  $\tilde{z}_{li}$  represents the cluster centre,  $\text{dis}_{EU \text{ improved}}(\tilde{z}_{li}, \hat{c}_{ji})$  represents the distance between the  $j$ th data point and the  $l$ th cluster centre;  $w_{lj}$  represents the semantic matrix;
- (7) if the number of iterations reaches  $\text{MaxStep}$  or satisfies  $|E_1 - E_2| < \epsilon$ , the iteration is terminated;  
Otherwise,  $O = O + 1$ ,  
return to step 5 and step 6
- (8) end;

ALGORITHM 1: Improved semantic similarity and relatedness-based K-means clustering algorithm.

experimental results of the SSE and Silhouette Coefficient are shown as follows.

Table 2 shows that as the number of  $K$  clusters increases, the SSE value of the improved K-means algorithm is significantly smaller than that of the traditional K-means algorithm. It shows that the clustering performance of the improved K-means algorithm is better than the traditional K-means algorithm. As shown in Figure 3, when the number of clusters ( $K$ ) equals 8, the improved K-means algorithm and the traditional K-means algorithm have an elbow (inflection point) within the SSE value. It shows that when the number of clusters is 8, performance clustering might be the best. It provides a reference for the value of  $K$  in K-means clustering.

As shown in Table 2, with the increasing number of clusters  $K$ , the Silhouette Coefficient value of the improved K-means algorithm is significantly higher than the traditional K-means algorithm. The higher value of the Silhouette Coefficient in the dataset indicates the better the clustering performance. Hence, it shows that the performance of the improved K-means algorithm is better than the traditional K-means algorithm. As shown in Figure 4, when the number of clusters ( $K$ ) equals 8, the improved K-means algorithm and the traditional K-means algorithm have relatively higher values. Hence, comprehensive analysis shows that the optimum value of  $K$  is 8. The red dotted lines in Figure 5 represent the Silhouette Coefficient of the traditional K-means algorithm and improved semantic K-means algorithm. The bar chart is the category of clusters. Most of the samples in a group have a higher Silhouette Coefficient value and are distributed near the red dotted line, representing a better clustering effect. On the contrary, if the sample points have a lower Silhouette Coefficient value and the distribution is scattered, the clustering effect is worse. Figure 5 shows that the Silhouette Coefficient value of the improved K-means algorithm is higher, and the sample distributes near the red

dotted line. Hence, the result indicates the performance of an improved K-means algorithm based on semantic similarity, and relatedness is better than the traditional K-means algorithm based on the Bag-of-Words model.

## 4. Experiment Results

**4.1. Data Collection.** The experimental data in this paper mainly collect the technological innovation information of 400 enterprises in Beijing and uses the document information as a text collection. The collected textual data mainly consist of the enterprise's primary status, the development of enterprise technological innovation activities, enterprise innovation projects, enterprise organizational structure, enterprise main products and services, enterprise profitability, etc. Table 3 briefly shows the details of the data collection result.

After data cleaning and selection, there are 867 valid texts, and the overall data size is about 20 M. The experimental operating environment is the Windows 10 system, 2.70 GHz core processor, 8.0 GB memory, and Python 3.6.2. After the preprocessing, including the custom dictionary, part-of-speech filtering, and stop words removing, the keywords vocabulary was obtained and shown in Figure 6. Then, map the keywords to the concept in domain ontology and get the semantic similarity matrix  $P$  and relatedness matrix  $Q$  by calculating the semantic similarity and relatedness.

**4.2. Results Analysis.** Figure 7 shows the result obtained by text clustering based on semantic similarity and relatedness. The most important 15 feature words in each cluster are selected to represent the topic based on the feature weight, as shown in Table 4. According to the topic reflected by each group, there are eight types of main factors affecting

TABLE 2: The performance comparison between improved K-means algorithm based on semantic and traditional K-means algorithm.

K Value	SSE		Silhouette coefficient	
	Traditional K-means algorithm based on bag-of-word model	Improved K-means algorithm based on semantic similarity and relatedness	Traditional K-means algorithm based on bag-of-word model	Improved K-means algorithm based on semantic similarity and relatedness
3	683.2779	672.4002	0.075280	0.363640
4	671.9479	577.2000	0.081084	0.355436
5	667.6426	532.9228	0.028152	0.352955
6	668.5414	519.1228	0.072469	0.355115
7	660.5026	461.7781	0.064319	0.328729
8	<b>645.9339</b>	<b>385.5498</b>	<b>0.075842</b>	<b>0.352197</b>
9	654.0991	392.1088	0.071101	0.362227
10	632.2624	373.0217	0.018480	0.344523

When the value of  $K$  equals 8, the improved  $K$ -means algorithm and the traditional  $K$ -means algorithm have a relatively higher value.

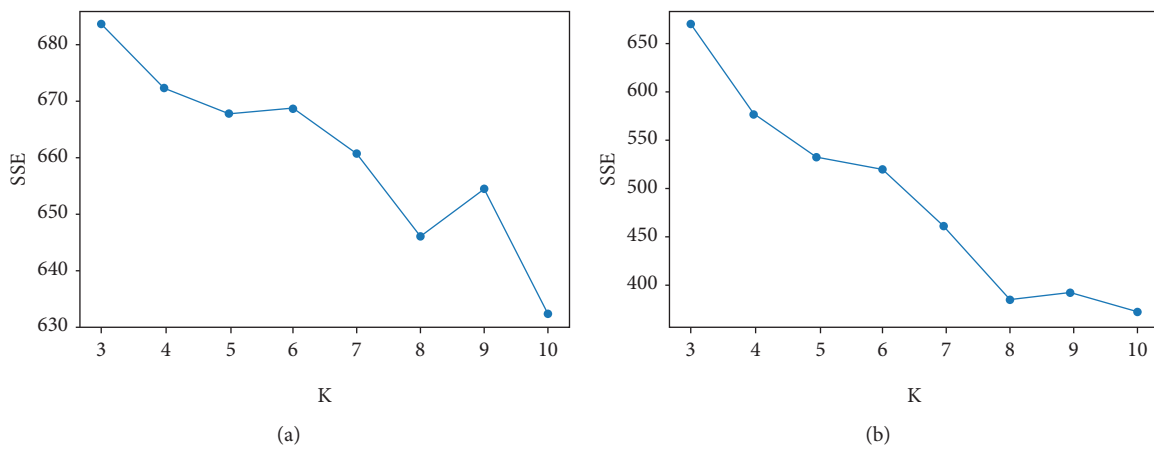


FIGURE 3: Statistical comparison of SSE. (a) Traditional K-means algorithm. (b) Semantic K-means algorithm.

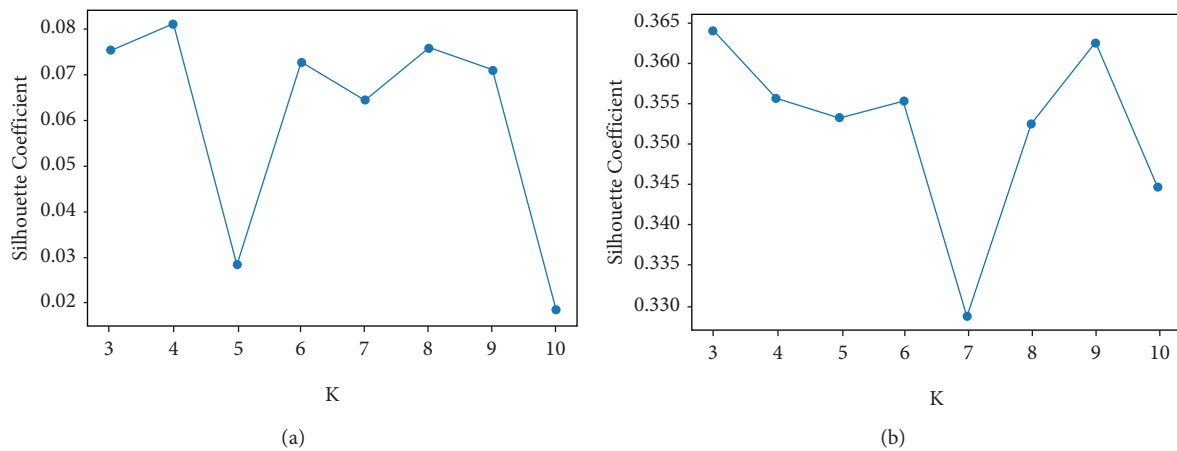


FIGURE 4: Statistical comparison of silhouette. (a) Traditional K-means algorithm. (b) Semantic K-means algorithm.

enterprise technological innovation. We combined the eighth cluster and the third cluster because they reflect the same theme. The analysis of the seven influencing factors about the technological innovation of enterprises is as follows.

Cluster 1: manufacturing capability. The main feature words of this cluster mainly focus on the new products, new processes, new materials, process technology, equipment level, etc. The content of these feature words is related to the manufacturing capabilities of products

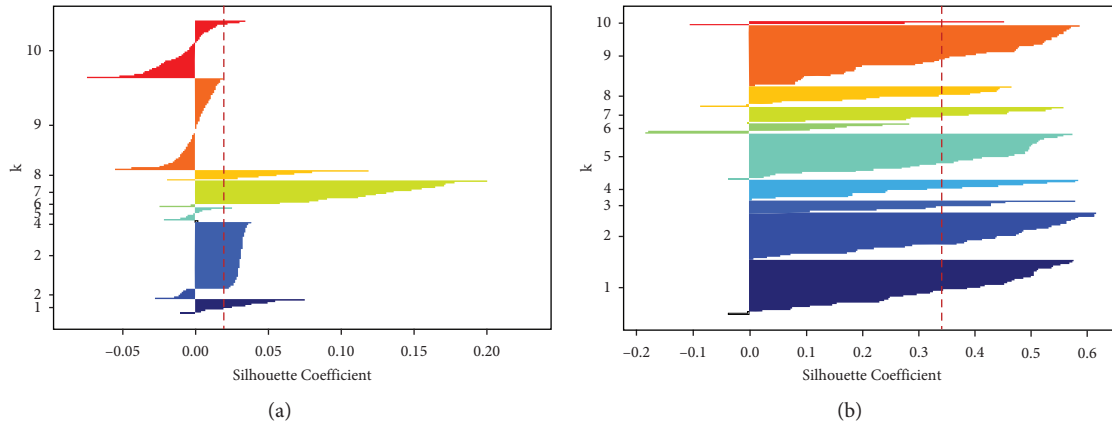


FIGURE 5: The histogram of silhouette coefficient. (a) Traditional K-means algorithm. (b) Semantic K-means algorithm.

TABLE 3: Summaries of data collection.

Area of focus	Format	Number
Enterprise profiles	.doc	388
Enterprise technical and financial reports	.xlsx	212
Enterprises products	.txt	131
Enterprises rewards	.pdf	136

No	doc_id	Partial content
1	0001	enterprise technical center, enterprise name, Airsys, Refrigeration Engineering, Technology, limited company, industry type, main business, development and production, self-produced product, innovation cooperation, technique fusion, strategic planning...
2	0002	Anton Oilfield Services, Technology, limited company, enterprise technical center, innovation system, construction situation, mechanism innovation, QHSE, Technological innovation activities, top talents, organization construction, management system, management system...
3	0003	Tianlong Tungsten-Molybdenum Technology, refractory materials, manufacturing service, equipment, high-technique, product research, manufacture factory, nonferrous metals, high and new technology, emerging industry...
4	0004	Antong construction, limited company, innovation trend, R&D team, core competitive advantage, university-industry cooperation, development tendency, engineering construction, resource integration...
5	0005	Airsys, technical center, data center, high availability, air conditioning equipment, refrigeration equipment, technical personnel, senior experts, total assets, equipment value...
6	0006	Austar Hansen, Packaging Technology, competitive advantage, corporate culture, innovation-driven, innovation management, enterprise development...
7	0007	Ankong, Technology Development, innovative product, solution, product seriation, industrial value chain, market demand, market research, core competitiveness, social benefit, brand influence...
8	0008	Orion Energy Technology Development, innovation trend, technological innovation system, operational condition, organizational construction, cooperation innovation, innovation project...
9	0009	Babcork Wilcox Beijing Company, innovation trend, competitive advantage, enterprise development, internal resources, innovation research, technical cooperation...
10	0010	Bestpowern electrical technology, main business, manufacturing, technical process, quality control, product research, R&D expenditure...

FIGURE 6: The keyword vocabulary after text preprocessing.

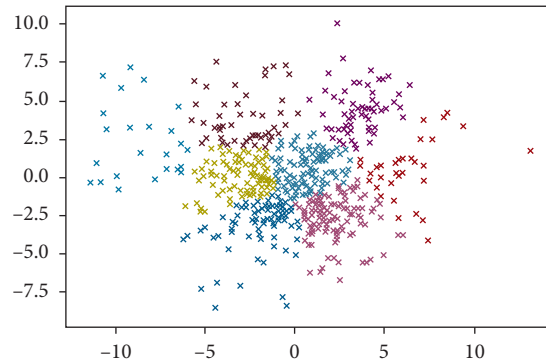


FIGURE 7: Visualization results of enterprise technology innovation text clustering when  $K=8$ .

TABLE 4: The text clustering result of enterprise’s technological innovation when  $K=8$ .

Cluster	Proportion of included samples (%)	Top 15 topic feature words
Cluster 1	16.1	New product, Internet application, new technology, construction technique, information technology, new technique, advanced level, composite material, design and development, quality control, process technology, technical process, industrialization, high-tech, equipment level, individuation
Cluster 2	16	Engineer, senior engineer, senior expert, R&D expenditure, R&D, total assets, employees’ number, equipment original value, expenditure on science and technology activities, main business proportion, bachelor degree or above, enterprise scale, asset-liability ratio ownership structure, technology introduction
Cluster 3	19.2	Mechanism innovation, organization and implementation, operating mechanism, rewards system, organizational construction, management system, organizational structure, performance reviews, post doctor, rules and regulations, excellent talents, organization, incentive mechanism, communication channel, guarantee mechanism
Cluster 4	9.7	Patent, industry-standard, gross profit on sale, main business product sales revenue, science and technology progress award, method number, number of patent applications, industrial output, industrial added value, platform, number of technology development projects, number of new product development projects, software copyright, utility model, design patent
Cluster 5	7.3	Competitive advantage, social benefit, brand influence, market competitiveness, core competence, reputation, diversification, business area, marketing, market-driven, market demand, development trend, market occupancy, market share, products sale
Cluster 6	5.7	Intellectual property, intellectual property protection, intellectual property management, independent intellectual property, intellectual property application, property rights transformation, transfer of property rights, license, patent warning, confidential agreement, trade secrets, technology protection, protection method, promote application, intellectual property layout
Cluster 7	21.7	Industry-university-research cooperation, industry-university-research, internal and external resources, integration, resource integration, core competence, strategic planning, research institutes, colleges and universities, R&D team, technical cooperation, technology exchange, overall planning, leader, technology fusion
Cluster 8	4.3	Technical staff, innovation mechanism, management mechanism, director of technology centre, manager, scientists, technology leader, positivity, talent team, innovation team, outstanding contribution, postgraduate, technical experts, talent incentive mechanism, youth technology leader

and processes. The influence of manufacturing capability on technological innovation of enterprises is mainly reflected in the capability to transform the research and development results into manufacturing. The word “equipment level” reflects the advanced manufacturing equipment, the phrases “Construction technique, Process technology, Technical process, High-tech” reflect the topic of process design capability, and the term “quality control” reflects the content of

product quality management. “Internet application, Information technology, Industrialization” reflect the theme of product innovation activities. Therefore, the cluster’s words with high feature weights reflect that manufacturing capability is essential for an enterprise’s technological innovation.

Cluster 2: innovation resources. The words “engineer,” “senior engineer,” and “R&D expenditure” in this cluster have a high proportion of feature weight. The



words “Engineer, Senior engineer, Senior expert, Employees number, Bachelor degree or above” reflect the quantity and quality of the R&D staff. The words “R&D expenditure, R&D, Total assets, Expenditure on science and technology activities, Main business proportion” reflect the financial investment on R&D. The word “Equipment original value” reflects the equipment investment of R&D. The phrases “Enterprise scale, Asset-liability ratio ownership structure, Technology introduction” represent the investment of enterprises in non-R&D including of enterprises own capability. The investment of innovation resources mainly refer to the quantity and quality of enterprise’s investment in technological innovation resources. It is reflected in the investment of staff, funds, and equipment in R&D. The investment of innovation resources is one of the influencing factors of enterprise technological innovation.

Cluster 3: mechanism innovation. The top feature words of this cluster can reflect directly that the topic is mainly about the mechanism innovation, which is an essential factor that affects enterprises’ technological innovation. The feature words such as “Rewards system, Performance review, Post-doctor, Excellent talents, Incentive mechanism” indicate that enterprises attach importance to the incentive mechanism of personnel. The words “Organization and implementation, Operating mechanism, Organizational construction, Management system, Organizational structure” reflect the influence of the organizational management mechanism on the technological innovation of enterprises. An effective innovation mechanism can stimulate talents and cooperate effectively. The eighth group’s theme is the same as cluster 3, and the sample proportion is only 4.3%. Hence, we merge the contents of cluster 3 and cluster 8.

Cluster 4: innovation output. The feature words in this cluster, such as “Patent, Industry standard, Gross profit on sale, Main business product sales revenue,” have higher feature weight. It indicates that this cluster mainly reflects the topic of innovation output. The words “Patent, Industry standard, Science and technology progress award, Method number, Number of patent applications, Number of technology development projects, Number of new product development projects, Software copyright, Utility model, Design patent” reflect the technological output of enterprise’s innovation. The words “Gross profit on sale, Main business product sales revenue, Industrial output, Industrial added value” reflect the innovation benefits. Therefore, from the distribution of feature words, the innovation output is also an important performance that affects the technological innovation of an enterprise.

Cluster 5: market innovation. The words “Competitive advantage, Social benefit, Market competitiveness” have high feature value in this cluster. The keywords tend to reflect market innovation, which means

product sales and promotion innovations can meet market demands. With the continuous expansion of market scale and the increase of market demand, the market-oriented product sales innovation model has become one of the important factors affecting the development of enterprise technological innovation.

Cluster 6: protection measures. “Intellectual property, Intellectual property protection, Intellectual property management, Independent intellectual property” are high feature value words in this cluster, reflecting the content of intellectual property protection measures. Technical knowledge protection can promote technology diffusion and ensure attracting foreign capital and technology introduction. Therefore, the protection measures for intellectual property and technology are conducive to promoting the technological innovation of enterprises and are an important influencing factor.

Cluster 7: innovation strategy. This cluster contains the most significant proportion of samples. The keywords “Industry-University-Research Cooperation, Industry-University-Research, Internal and external resources, Integration, Resource Integration” reflect the topic of innovation strategy. It refers to the integration and arrangement of internal and external innovation resources and technologies based on the enterprise’s overall strategy with enterprise operation. The phrases “Industry-University-Research Cooperation, Industry-University-Research, Research institutes, Colleges and universities, R&D team” represent the content of the joint innovation strategy of enterprises and industry, university, and research. The phrases “Internal and external resources, Resource Integration” reflect the integration and allocation of internal and external resources. The terms “Technical cooperation, Technology Exchange, Technology fusion” reflect the strategic plan of enterprises for technology integration and innovation. The phrases “Core competence, Strategic planning, Overall planning, Leader” represent the innovation capability and strategy of the enterprise leader or decision-maker. Therefore, the topic of this cluster is innovation strategy, and the correctness of the innovation strategy also has an important impact on enterprise technological innovation.

Figure 8 shows the framework of the influencing factors for enterprises’ technological innovation. There are 7 types of impacts factors: manufacturing capability, innovation resource, mechanism innovation, innovation output, market innovation, protection measures, and innovation strategy. Through analysis of the linking feature words of each cluster, the meso-level concept can be concluded, as shown in Figure 8.

## 5. Discussion

*5.1. Policy Suggestions.* It can be seen from the framework of influencing factors of enterprise’s technological innovation, and the four clusters results that are particularly prominent include the “Protection measure,” “Innovation strategy,”

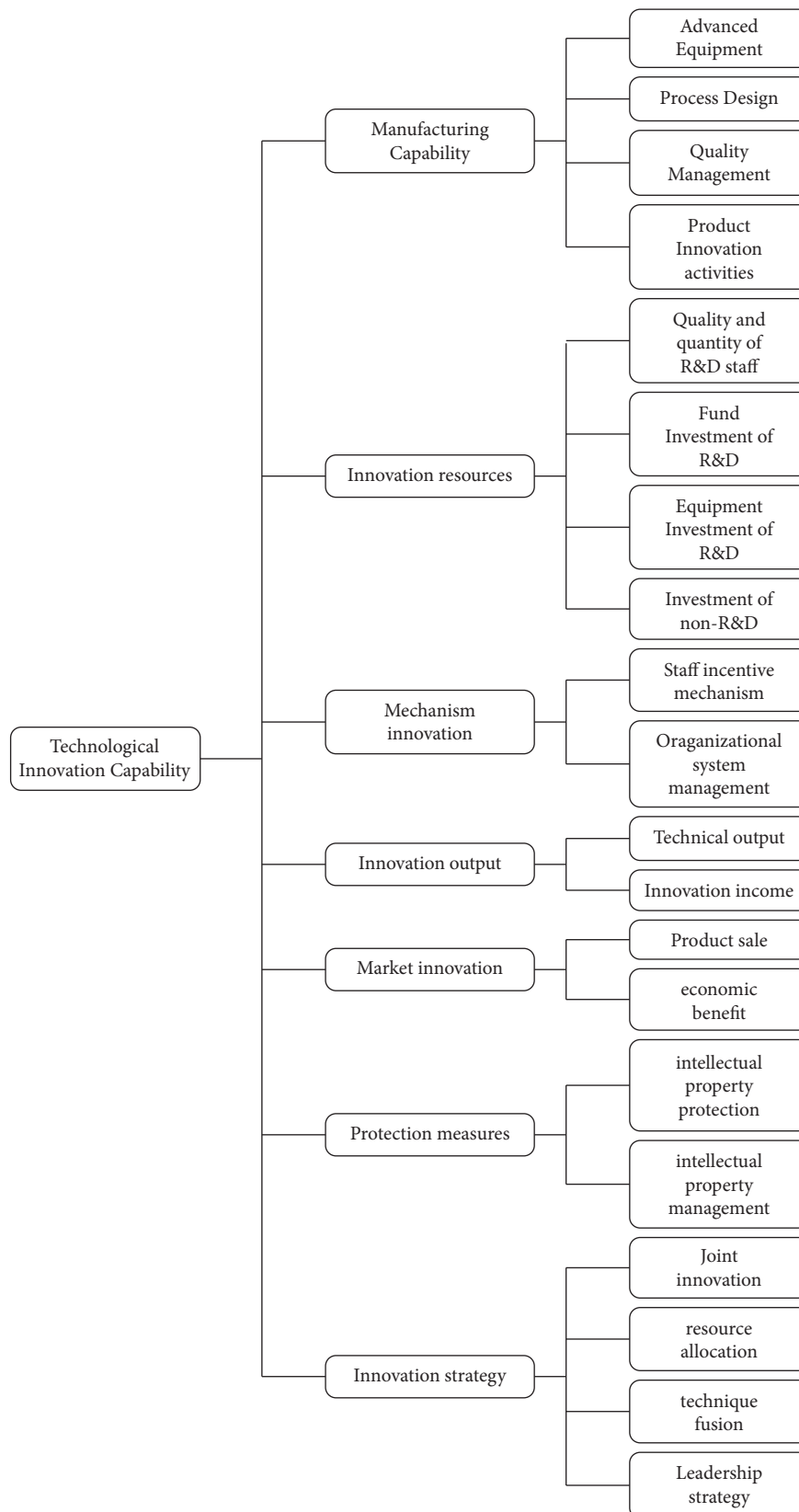


FIGURE 8: The framework of influencing factors of enterprise's technological innovation.

“Market innovation,” and “Mechanism innovation” except the “Manufacturing capability,” “Innovation resources,” and “Innovation output,” which are included in most of the researches. The four aspects are relatively new in the field of enterprise’s technological innovation. Thus, the paper provides policy suggestions from these four aspects for the enterprise’s technological innovation.

The innovation capability of intellectual property is an important aspect to measure the innovation output of enterprises. In the current fierce competition environment, enterprises are rushing to develop scientific and technological achievements through various channels for survival and development, and seek legal protection by applying for patents. However, it is not uncommon for enterprises to suffer heavy losses because their competitors registered the patents and trademarks in advance. Therefore, enterprises should strengthen their awareness of intellectual property protection, set up special intellectual property departments, and pay attention to cultivating professional intellectual property talents.

The innovation strategy is a plan and methodology for enterprises to develop new products and services in the future. It aligns the development of innovations with future corporate goals, which requires formulation based on the external environment and internal conditions. Thus, the innovation strategy directly reflects leadership decision-making. The weak sense of innovation or lack of innovative decision-making power for senior decision-makers will bring risks to the enterprise. The enterprises should plan and layout the innovative strategic cooperation in advance. In addition, it strengthens the collaboration between schools and industry and establishes strategic partners, which is helpful for the enterprise’s technological innovation.

Market innovation is mainly reflected in marketing and management capabilities. From the formation of initial products to the market introduction, the whole process is inseparable from professional marketers and marketing strategies. Therefore, to strengthen the market development, enterprises need to expand sales channels, explore new sales models, and improve the sales platforms. Moreover, enterprises should build their own sales team, introduce professional marketing personnel, and strengthen the training and management of marketing.

Introduce innovative management talents and improve the enterprise innovation mechanism. The enterprises need to construct a suitable organizational framework and formulate the innovation management system, including a talent introduction mechanism, talent training mechanism, innovation incentive system, reasonable innovation evaluation system, innovation achievements protection mechanism, etc. It can mobilize the enthusiasm of corporate managers and employees by improving the incentive mechanism, including interest incentives, competence incentives, power incentives, and responsibility incentives. The interest incentives include salary, welfare, bonus, etc. Competence incentives include training, competitive employment, etc. The power incentives mainly refer to promotion, and responsibility incentives mainly refer to a reasonable assessment system. Mobilizing the enthusiasm

and creativity of employees can effectively enhance the enterprise’s technological innovation capabilities.

*5.2. Implication for Theory and Practice.* This paper enriches and deepens the theoretical research on evaluating enterprise’s technological innovation capability from the theory aspect. The paper first extracts prominent factors that affect enterprise technological innovation based on the collected textual data. It provides a reference for the construction method of the enterprise’s technological innovation evaluation index system. Moreover, the process of enterprise’s technological innovation is dynamic and continuous. Traditional assessment methods for enterprise technology innovation lack automatic processing ability and cannot meet the needs of large-scale, high-quality, and in-depth knowledge acquisition. As a high-tech knowledge processing technology, the text mining method has significant advantages in the intelligent analysis of enterprise technological innovation capabilities. It can reflect the objective level of enterprise’s technological innovation under the data-driven. This paper proposes a method for discovering enterprise technological innovation knowledge based on semantic mining technology. It is helpful to explore the potential factors of influencing the enterprise technological innovation by revealing the inherent complex associations of enterprise’s textual data and extracting valuable patterns and knowledge. To our knowledge, this research is the first attempt to apply the semantic text clustering algorithm in enterprise’s technological innovation. In addition, this paper analyzed the development status and trends of the enterprise’s technological innovation capabilities based on the collected textual data from 400 enterprises in Beijing. It helps to promote technological innovation of the enterprises by smart analysis under the data-driven. Thus, the potential benefits of the proposed model can help drive and facilitate the enterprise’s technical innovation capability.

## 6. Conclusions

This paper utilizes a data-driven intelligent mining method based on a semantic conceptual model to analyze the influencing factors of enterprise technology innovation. Some scholars believe that technological innovation’s evaluation system construction method is a questionnaire design and survey analysis. However, it inevitably involves human factors and the subjective judgments of experts. This study proposed a systematic process for evaluating the enterprise’s technological innovation based on sizeable textual data. Computer information processing analyzes the evaluation factors and indicators that affect enterprise technology innovation from objective data. Furthermore, the traditional text clustering algorithm based on the Bag-of-words model ignores the semantic relationship between concepts, resulting in an unsatisfactory clustering effect. Therefore, this paper proposed an improved semantic concept-based clustering algorithm to analyze enterprise’s collected textual data. The performance of the improved K-means clustering method based on semantic similarity

and relatedness is superior to the traditional K-means clustering method. The proposed method realized the clustering of critical factors in the field of enterprises technological innovation, and the analysis of experimental results can obtain the seven key factors that affect the technological innovation of enterprises.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The author declares that there are no conflicts of interest regarding the publication of this paper.

## Acknowledgments

This research was supported by the Youth Research Fund Project of Beijing Wuzi University (2022XJQN19).

## References

- [1] X. Wang, X. Yang, X. Wang, M. Xia, and J. Wang, "Evaluating the competitiveness of enterprise's technology based on LDA topic model," *Technology Analysis & Strategic Management*, vol. 32, no. 2, pp. 208–222, 2020.
- [2] G. Richards, W. Yeoh, A. Y. L. Chong, and A. Popovič, "Business intelligence effectiveness and corporate performance management: an empirical analysis," *Journal of Computer Information Systems*, vol. 59, no. 2, pp. 188–196, 2019.
- [3] Q. Zhang, S. Liu, D. Gong, and Q. Tu, "A latent-dirichlet-allocation based extension for domain ontology of enterprise's technological innovation," *International Journal of Computers, Communications & Control*, vol. 14, no. 1, pp. 107–123, 2019.
- [4] J. Wang and Y. J. Chen, "A novelty detection patent mining approach for analyzing technological opportunities," *Advanced Engineering Informatics*, vol. 42, Article ID 100941, 2019.
- [5] L. Feng, Y. Niu, Z. Liu, J. Wang, and K. Zhang, "Discovering technology opportunity by keyword-based patent analysis: a hybrid approach of morphology analysis and USIT," *Sustainability*, vol. 12, no. 1, p. 136, 2020.
- [6] J. Jeon, C. Lee, and Y. Park, "How to use patent information to search potential technology partners in open innovation," *Journal of Intellectual Property Rights*, vol. 16, no. 5, pp. 385–393, 2011.
- [7] X. Zhou, L. Huang, Y. Zhang, and M. Yu, "A hybrid approach to detecting technological recombination based on text mining and patent network analysis," *Scientometrics*, vol. 121, no. 2, pp. 699–737, 2019.
- [8] H. J. Kim, T. San Kim, and S. Y. Sohn, "Recommendation of startups as technology cooperation candidates from the perspectives of similarity and potential: a deep learning approach," *Decision Support Systems*, vol. 130113229 pages, 2020.
- [9] S. Choi, H. Kim, J. Yoon, K. Kim, and J. Y. Lee, "An SAO-based text-mining approach for technology roadmapping using patent information," *R & D Management*, vol. 43, no. 1, pp. 52–74, 2013.
- [10] M.-Y. Wang, S.-C. Fang, and Y.-H. Chang, "Exploring technological opportunities by mining the gaps between science and technology: microalgal biofuels," *Technological Forecasting and Social Change*, vol. 92, pp. 182–195, 2015.
- [11] H. Park, J. J. Ree, and K. Kim, "Identification of promising patents for technology transfers using TRIZ evolution trends," *Expert Systems with Applications*, vol. 40, no. 2, pp. 736–743, 2013.
- [12] C. V. Trappey, H.-Y. Wu, F. Taghaboni-Dutta, and A. J. C. Trappey, "Using patent data for technology forecasting: China RFID patent analysis," *Advanced Engineering Informatics*, vol. 25, no. 1, pp. 53–64, 2011.
- [13] K. Song, K. S. Kim, and S. Lee, "Discovering new technology opportunities based on patents: text-mining and F-term analysis," *Technovation*, vol. 60-61, pp. 1–14, 2017.
- [14] V. Kayser and K. Blind, "Extending the knowledge base of foresight: the contribution of text mining," *Technological Forecasting and Social Change*, vol. 116, pp. 208–215, 2017.
- [15] J. S. Du, S. F. Peng, and J. S. Peng, "Research on technology innovation risk evaluation of high-tech enterprises based on fuzzy evaluation," *Journal of Intelligent and Fuzzy Systems*, vol. 38, no. 6, pp. 6805–6814, 2020.
- [16] A. Suder and C. Kahraman, "Multicriteria analysis of technological innovation investments using fuzzy sets," *Technological and Economic Development of Economy*, vol. 22, no. 2, pp. 235–253, 2016.
- [17] C. Feng and R. Ma, "Identification of the factors that influence service innovation in manufacturing enterprises by using the fuzzy DEMATEL method," *Journal of Cleaner Production*, vol. 253, Article ID 120002, 2020.
- [18] R. Mu, S. Xiao, and Y. Zhang, "Assessing innovation capability of small and medium-sized enterprises based on analytic hierarchy process," in *Proceedings of the 2018 International Conference on Education Technology, Economic Management and Social Sciences (ETEMSS 2018)*, vol. 1, pp. 383–387, Dalian, China, June 2018.
- [19] X. Pan, C. Han, X. Lu, Z. Jiao, and Y. Ming, "Green innovation ability evaluation of manufacturing enterprises based on AHP–OVP model," *Annals of Operations Research*, vol. 290, no. 1, pp. 409–419, 2020.
- [20] J. Liu, Y. Yin, and S. Yan, "Research on clean energy power generation-energy storage-energy using virtual enterprise risk assessment based on fuzzy analytic hierarchy process in China," *Journal of Cleaner Production*, vol. 236, Article ID 117471, 2019.
- [21] Y. Wang, J. Pan, R. Pei, B. Yi, and G. Yang, "Assessing the technological innovation efficiency of China's high-tech industries with a two-stage network DEA approach," *Socio-Economic Planning Sciences*, vol. 71, Article ID 100810, 2020.
- [22] X. Ma, Z. H. Liu, Y. F. Gao, and L. Na, "Innovation efficiency evaluation of listed companies based on the DEA method," *Procedia Computer Science*, vol. 174, pp. 382–386, 2020.
- [23] H. Li, H. He, J. Shan, and J. Cai, "Innovation efficiency of semiconductor industry in China: a new framework based on

- generalized three-stage DEA analysis,” *Socio-Economic Planning Sciences*, vol. 66, pp. 136–148, 2019.
- [24] T. Shang, “Evaluation model and index system for enterprise strategic management capability based on neural network algorithms,” *International Journal of Engineering Intelligent Systems*, vol. 27, no. 2, pp. 79–83, 2019.
- [25] Z. Zhen and Y. Yao, “Lean production and technological innovation in manufacturing industry based on SVM algorithms and data mining technology,” *Journal of Intelligent and Fuzzy Systems*, vol. 37, no. 5, pp. 6377–6388, 2019.

## Retraction

# Retracted: An Individual Driving Behavior Portrait Approach for Professional Driver of HDVs with Naturalistic Driving Data

### Computational Intelligence and Neuroscience

Received 3 October 2023; Accepted 3 October 2023; Published 4 October 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] Y. He, S. Yang, X. Zhou, and X. Lu, "An Individual Driving Behavior Portrait Approach for Professional Driver of HDVs with Naturalistic Driving Data," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3970571, 14 pages, 2022.



## Research Article

# An Individual Driving Behavior Portrait Approach for Professional Driver of HDVs with Naturalistic Driving Data

Yi He <sup>1</sup>, Shuo Yang<sup>1,2</sup>, Xiao Zhou <sup>3</sup>, and Xiao-Yun Lu<sup>4</sup>

<sup>1</sup>Intelligent Transport Systems Research Center, Wuhan University of Technology, Wuhan 430063, China

<sup>2</sup>CATARC (Tianjin) Automotive Information Consulting Co., Ltd., Beijing 100011, China

<sup>3</sup>School of Mechanical and Electronic Engineering, Wuhan University of Technology, Wuhan 430063, China

<sup>4</sup>California PATH, University of California, Berkeley, CA 94804-4684, USA

Correspondence should be addressed to Xiao Zhou; [zhouxiao@whut.edu.cn](mailto:zhouxiao@whut.edu.cn)

Received 30 October 2021; Accepted 30 November 2021; Published 22 January 2022

Academic Editor: Daqing Gong

Copyright © 2022 Yi He et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

More than 50% major road accidents are caused by risk driving behaviors from professional drivers of Heavy Duty Vehicles (HDVs). The quantitative estimation of driving performance and driving behaviors portrait for professional drivers is helpful to measure the driver's driving risk and inherent driving style. Previous studies have attempted to evaluate risk driving behavior, but most of them rely on high-demand vehicle and driving data. However, few studies can dig into the causes and correlations behind individual driving behavior and quantify the driving behaviors portrait for professional driver based on long-term naturalistic driving. In this study, the data is from On-Board Unit (OBU) devices mounted in the HDVs in China. Based on the driving behavior pattern diagram and the frequency and ranking of drivers' typical driving patterns, a driving behavior portrait approach is proposed by comprehensively considering the vehicle safety, driving comfort, and fuel economy indicators. The similarities and differences of different drivers' driving behaviors are quantitatively analyzed. The high precision and sampling frequency data from vehicles are used to verify the proposed approach. The results demonstrated that the driving behavior portrait approach can digitally describe the individual driving behaviors styles and identify the potential driving behaviors with long-term naturalistic driving data. The development of this approach can help quantitatively evaluate the individual characteristic of risk driving behaviors to prevent road accidents.

## 1. Introduction

More than 50% major road accidents are caused by risk driving behaviors from professional drivers of HDVs [1]. Due to drivers' personal characteristic and driving style, most of the road accidents are caused by risk driving behaviors [2]. Quantifying analysis drivers' driving style and identifying driving behavior portrait is helpful to prevent road accidents.

In recent years, many studies have attempted to evaluate risk driving behavior, but most of them rely on high-demand vehicle and driving data, which is not conducive to large-scale application, and most of the previous studies focus on driving safety levels or identify the driving behaviors based on real time or offline with short-term data analysis ([3–12]; Yan et al., 2019). However, few studies can dig into the

causes and correlations behind individual driving behavior and quantify the driving behaviors portrait for professional driver based on long-term naturalistic driving.

Various driving behaviors can be seen as the responses to complex driving conditions. Due to the personal characteristics and driving skills of different drivers, their driving behaviors are also different. Some drivers have aggressive driving styles, fast acceleration, and braking with high risk driving, while others are more cautious with less risk. Therefore, the quantitative estimation of driving performance and driving behaviors portrait is helpful to measure the driver's driving risk and inherent driving style, so as to prevent traffic accidents.

With the development of autonomous and connected vehicle, various sensors can be installed on the vehicle to collect real time data with naturalistic driving. However, the

positioning data is the most widely used and has high precision with large-scale application in real traffic until now.

In this paper, the data is from OBD devices mounted in the HDVs in China, and high precision and sampling frequency data from vehicles are used to verify the proposed approach. Based on the driving behavior pattern diagram and the frequency and ranking of drivers' typical driving patterns, a driving behavior portrait approach is proposed by comprehensive considering the vehicle safety, driving comfort, and fuel economy indicators, to quantitatively evaluate the personalized driving behavior of different drivers and identify potential risk driving behaviors, so as to improve the assessment accuracy of the risk model and reduce the accident risk.

## 2. Literature Review

The majority of road accidents are caused by risk driving behaviors of professional drivers with HDVs. Therefore, scholars from all parties have carried out assessment studies on drivers' driving safety and driving risk, mainly from the following three aspects.

*2.1. Risk Driving Behavior Identification Based on Statistical Methods.* Some risk driving behavior studies are related by applying statistical method or machine learning algorithm to classify and identify risk driving. Constantinescu et al. modeled driving styles based on positioning data and divided drivers into different driving styles by cluster analysis and principal component analysis [8]. Suzdaleva et al. studied the stochastic modeling of driving style and its online estimation during driving based on Bayesian theory, and their research used recursive algorithm to estimate hybrid parameters and actual driving style [7]. Martinez et al. introduced a vehicle driving style feature and recognition correction algorithm based on current and future development trends with machine learning methods [9]. Qi et al. studied two thematic models: mLDA and mHLDA, which can discover the information of differentiated driving styles with hidden structures from the data of real driving behaviors and can identify the different driving styles [10]. Zhu et al. proposed an unsupervised clustering algorithm that can reflect different driving styles [13]. Han et al. proposed a method of driver behavior uncertainty identification based on statistics [11]. Wang et al. used semi-supervised support vector machines to classify driving styles. In order to solve the problem that supervised learning method requires a large amount of marked training data and it is time-consuming to manually label a large amount of driving data, a semisupervised support vector machine method is used in this paper to classify drivers into aggressive driving style and conventional driving style according to several marked data points [6]. Siordia et al. proposed an autonomous driving risk rating system. The system is based on a driving simulator that uses driver, road, and vehicle information to train five different data mining algorithms to predict the level of driving risk [14].

*2.2. Risk Driving Behavior Identification Based on Safety-Related Driving Events Detection.* Another commonly used method for classifying driving risks is to detect safety-related driving events, such as sharp acceleration, sharp deceleration, and sharp turns. In Johnson et al.'s study, a dynamic time structuring algorithm (DTW) was used to detect the driving styles with data from smartphone sensors [15]. In this case, they need some predefined templates for the detected driving events. Itkonen has made a trade-off between the compact but erratic behavior of driving and the loose and smooth following behavior of driving. They call it "tight" and "smooth" driving and suggest that this trade-off can be used as a potential factor to evaluate potential driving styles [16]. In these studies, the use of predefined driving events is essential, but the thresholds used to define these driving events are not consistent. For example, in Paefgen et al.'s study, the threshold value used for rapid acceleration and deceleration is  $\pm 0.1$  g [17]. Fazeen uses  $\pm 0.3$  g [3]. Wang et al. proposed a self-designed driving style scale, which is suitable for evaluating the driving style of drivers with Chinese national conditions [18]. Bagdadi proposed a method based on critical jerk, to identify safety critical braking events during car driving [19]. Chen et al. proposed a method to predict risk driving events and defined risk driving events by using the risk attribute diagram (ARM) [4].

*2.3. Risk Driving Behavior Identification Based on Risk Modeling.* Some scholars have also modeled the safety of drivers' driving behavior. Ellison et al. assessed the risk driving behavior by modeling driver behavior. The frequency and size of overspeed event, aggressive acceleration event, and aggressive braking event were used as the behavior measurement indexes, and their spatial and temporal characteristics were used as the influencing factors. Based on these parameters, the behavioral characteristics of drivers were established and the total score representing driving risk was given [5]. Bejani et al. developed the system to consider the impact of traffic levels and vehicle types on driving style assessments. The system consists of three subsystems, which use smart phones for calibration, driving action classification, and driving style evaluation [20]. Hong et al. used sensors such as Android smartphones, On-Board Diagnostics (OBD), and Inertial Measurement Units (IMU). By collecting driving behavior characteristics such as maximum speed and acceleration, average value, and standard difference, the relationship between driving characteristics and driving style was modeled by using naive Bayes classifier [21]. Jachimczyk et al. proposed a method to objectively evaluate driving style based on eight indexes related to speed, acceleration, engine speed, and driving time. This method is an embedded system based on the conceptual design of the Internet of things to evaluate three driving style standards: safety, economy, and comfort [22]. Meseguer et al. designed an Android application that can monitor physiological data from drivers and diagnostic data from vehicles in real time to study the correlation between heart rate and driving style [23]. Chu et al. proposed an improved curve speed model that takes into account driving style, vehicle, and road

factors. On the basis of vehicle-road interaction model, driver behavior factor is introduced to quantify the driving style selected for the speed of curve [12, 24]. Deng et al., in order to identify driving styles, reasonably modeled three driving styles (aggressive, moderate, and mild) by means of hidden Markov model according to the driver's braking characteristics and realized efficient analysis of driving styles [25]. An early warning system has been proposed for risk driving based on statistical modeling. They used a semi-supervised learning method to grade the driving risk and evaluate the safety of the driving time series data [26].

In summary, most of the risk driving behaviors studies focus on the classification of driving styles or the scoring of risk driving behaviors. The purpose of these previous studies is mostly to classify drivers into different driving safety levels, such as safe, normal, or dangerous, and rarely pay attention to individual driving behavior characteristics. In addition, the use of predefined risk driving behaviors models or events is essential, but the thresholds used to define these driving events are not consistent. However, a few studies focusing on the risk driving behavior model for professional driver of HDVs and a fixed risk driving threshold may not fit in all drivers due to various driving styles from individual drivers and understanding of the driving skills and experience.

In recent years, Internet portrait technology is widely used to quantify the characteristics of users. A driving habit graph (DHG) was proposed to model driving behaviors graphical model [4], which can quantify the behaviors parameters in naturalistic driving data and visual presentation of driving behaviors. Inspired by that study, this paper proposes an individual driving behavior portrait by comprehensively considering the vehicle safety, driving comfort, and fuel economy indicators to quantify the potential risk driving behaviors and driving styles.

### 3. Methodology

**3.1. Driving Behavior Pattern Diagram.** Driving behavior is the driver's response to different driving situations, and driving behavior patterns reflect these continuous driving behaviors over time; in this case, each driver has individual typical driving mode. We need to identify these driving behavior patterns and extract the representative ones to judge the driver's driving style based on the driving patterns of different drivers to evaluate the driver's driving safety level. In the process of method exploration, we use the Symbol Aggregation Approximation (SAX) method to convert the original driving data into the corresponding driving acceleration level code to reduce the complexity of original driving data [27]. SAX algorithm was proposed by Lin et al. in 2003 [28]. Based on the Piecewise Aggregation Approximation (PAA) method, this method retains its advantages of simplicity and low complexity and provides high sensitivity and selectivity in the process of range query. In addition, the symbolic feature representation provides a favorable condition for the existing rich data structures and string processing algorithms. The SAX process consists of two steps. First, convert the original data into a segmented

aggregation approximation (PAA) representation. Second, convert PAA data to the corresponding string according to the location of the breakpoint. In this study, the driving behavior data was ultimately converted into a numeric code rather than a string. This research includes four steps: data standardization, data grading and coding, typical driving behavior pattern extraction, and driving behavior pattern graph generation.

**3.1.1. Data Standardization.** In this study, the longitudinal acceleration change of the driver is used as the standard to reflect the driving behavior. Longitudinal acceleration is chosen as the standard because acceleration is a typical index closely related to driving safety, and the change of acceleration also represents the change of speed, which is another important index of driving safety.

The driving acceleration data of a certain driver is standardized by Z-standardization method, so that its mean value is 0 and the standard deviation is 1. The purpose of this is to make the offset and amplitude consistent so that it has the same dimension as other driver data.

**3.1.2. Data Grading and Coding.** The grading and encoding of data require the use of SAX method. The SAX method first converts the original data into PAA representation. The PAA data is then converted to the appropriate string based on the location of the breakpoint.

This study divided the data into 7 levels. Theoretically, the SAX method can divide the data into up to 20 levels and provide corresponding thresholds. Based on data analysis, after testing the classification of 3, 5, 7, 9, and 11 grades, we found that the classification of 7 grades had the best effect. Only the odd number level is tested because the median level always represents a neutral behavior. In this study, the median represents the driving behavior of the vehicle approaching uniform cruising speed. Since the mean of the normalized data is 0, level 3 represents neutral behavior. Taking the acceleration data as an example, level 3 represents the acceleration value close to  $0 \text{ m/s}^2$ , that is, uniform speed cruising. Levels 0, 1, and 2 represent negative acceleration, or deceleration. The amplitude range of data values for each level increases as the level decreases. As the number of levels decreases, the deceleration behavior becomes more obvious. Levels 4, 5, and 6 represent positive accelerations, or accelerations. The amplitude range of data values for each level increases with the level. As the rank value increases, the accelerating behavior becomes more obvious.

**3.1.3. Typical Driving Behavior Pattern Extraction.** After the data is graded and coded, the driving mode that exists in it needs to be detected. Driving is a continuous process in which the vehicle is controlled by the driver's constant adjustment to achieve the desired driving condition. Limited by people's physiological and psychological factors, the reaction speed of the driver to the stimulus is limited. Under general conditions, the reaction time for complex selective reactions is 1~3 s, and the reaction time for complex



judgment and recognition is 3~5 s on average. Studies have shown that three seconds is a window of time to measure driving behavior, during which the driver can adjust the vehicle to the desired state. Thus, the driving mode is defined as a change in driving behavior (speed, acceleration, steering, etc.) every three seconds. To make sure the timing is right, this article also tested using 4, 5, and 6 seconds. The results show that if these intervals are used as the detection time, a large number of driving mode types are generated, and the graph is not as clear as the result of 3 seconds. Starting from the starting point of the data, a moving window of window size 3 s is used to detect the existing driving mode. Due to the differences in driving skills and personal characteristics, each driver has his own unique driving habits. According to all the driving behavior patterns detected by different drivers, the driving behavior patterns with high frequency are used to represent the driving habits of drivers. For each driver, more frequent driving behavior patterns, that is, typical driving behavior patterns, were extracted. Driving patterns with the same sequence of combinations naturally converge into a group and count the number of patterns contained in each group. Then, the incidence of each driving behavior pattern was calculated. For example, if driving mode “123” occurs “ $m$ ” times in a continuous acceleration data set of length “ $n$ ,” the incidence of driving behavior mode “123” is “ $m/(n-2)$ ”. For each driver, the 30 most frequent driving behavior patterns were extracted and regarded as their typical driving behavior patterns. On the premise of collecting enough driving data to cover various external factors such as traffic conditions and road types, these typical driving behavior patterns can describe drivers’ internal driving habits. In addition, these typical driving behavior patterns are independent of external driving conditions and reflect drivers’ inherent driving preferences.

**3.1.4. Driving Behavior Pattern Graph Generation.** The typical driving patterns were plotted in descending order of frequency to obtain the driving behavior pattern of individual drivers. Driving behavior data can be transformed into a graph through the four steps described above. The typical driving behavior pattern reflects the driver’s personalized driving behavior, but also includes the driver’s driving styles and hidden driving risks. In this study, the driving behavior patterns of drivers were plotted based on the acceleration data. For all driving behaviors, the first 30 driving patterns are extracted as the typical driving patterns. These 30 patterns are enough to reflect the driver’s driving behavior characteristics. In addition, considering that acceleration and deceleration are two important and independent operations, the typical driving patterns of acceleration and deceleration are extracted, respectively, to illustrate the driving behavior characteristics of drivers and distinguish the differences between drivers’ driving styles.

The abscissa in Figure 1 represents the first 30 typical driving behavior patterns that appear from high to low frequency after the extraction of typical driving patterns. The ordinate represents the seven acceleration levels of 0–6 after data classification and coding. That is, the combination of

every three connected line points in the figure is a typical driving behavior pattern, and the corresponding abscissa is the occurrence frequency ranking of the behavior pattern, while the ordinate represents the corresponding acceleration level change of the driving behavior pattern within three seconds. Acceleration level 3 represents the neutral behavior, which is close to the acceleration value of 0 m/s<sup>2</sup>, that is, the average speed of travel. Acceleration levels 0, 1, and 2 represent the negative acceleration value, that is, the deceleration behavior. The amplitude range of data values for each level increases as the level decreases. As the number of levels decreases, the deceleration behavior becomes more obvious. The acceleration levels of 4, 5, and 6 represent the positive acceleration value, that is, the acceleration behavior. The amplitude range of data values for each level increases with the level. As the rank value increases, the accelerating behavior becomes more obvious. Taking the first typical driving behavior pattern “333” as an example, the frequency of this pattern is the highest. During the duration of this behavior pattern, the driver’s acceleration level changes to “3-3-3.” The second driving behavior pattern, “343,” was the second most frequent. During the 3 seconds of this behavior pattern, the driver’s acceleration level changed to “3-4-3.” The driver’s driving behavior pattern diagram can clearly present the composition and sequence of the driver’s typical driving behavior pattern.

**3.2. Modeling Driving Behavior Risk.** The driving behavior pattern graph is the visual representation and qualitative description of the driver’s driving pattern. However, the frequency and ranking of typical driving patterns and the gradation of the patterns themselves offer possibilities for quantitative analysis. Therefore, taking acceleration data and typical driving patterns as influencing factors, this paper proposes a personalized driving behavior evaluation model from three aspects of vehicle safety, driving comfort, and fuel economy to comprehensively evaluate vehicle driving safety.

$$\text{SCORE} = 0.7 * \text{score}_{\text{safe}} + 0.2 * \text{score}_{\text{comfort}} + 0.1 * \text{score}_{\text{economy}} \quad (1)$$

As shown in formula (1), SCORE is the comprehensive evaluation index of driving behavior, which consists of three parts. The first part is the safety evaluation index based on the driver’s typical driving behavior, which reflects the driver’s driving safety degree and the severity of acceleration and deceleration. The second part is the evaluation and analysis of vehicle comfort based on acceleration data, which reflects the degree of comfort in driving behavior to some extent. The third part is the evaluation index of fuel economy based on acceleration, which reflects the degree of fuel consumption. Combined with literature, considering the operation of the vehicle in the process of driving safety is most important, followed by the comfort in the process of transportation, and finally to fuel economy, giving three sections of the score to the corresponding proportion of weight (i.e., safety weight with 70%, comfort with 20%, and the economy with 10%), comprehensive comparing to the driver’s driving behavior evaluation.

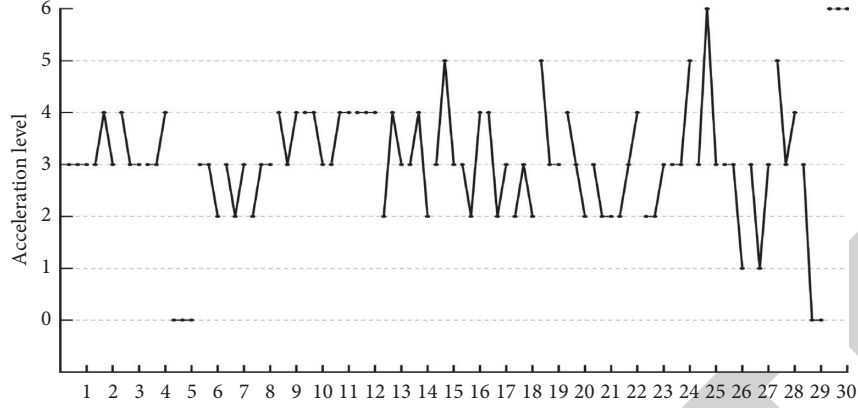


FIGURE 1: Driving behavior pattern diagram.

$$\text{score}_{\text{safe}} = \sum_{i=1}^{30} \text{freq}_i * \{ \text{abs}[\text{mean}(\text{CODE}) - 3] + \text{std}(\text{CODE}) \} \quad (2)$$

As shown in formula (2), in the first part,  $i$  is the order of the typical driving mode, and  $\text{freq}_i$  is the proportion of the typical driving pattern in all driving patterns. CODE is the code for a typical driving pattern, such as "234." Mean and std were used to calculate the mean and standard deviation of driving patterns, respectively. For example, the mean of pattern "234" is  $(2 + 3 + 4)/3 = 3$ , and the standard deviation is 1. Abs stands for absolute value. This evaluation method can reflect the degree of safety of driving behavior, because it takes into account the amplitude and variation of acceleration value by using the mean value and standard deviation. The amplitude of driving acceleration and the change of acceleration are important indexes, so the mean value and standard deviation are added. In  $\text{abs}[\text{means}(\text{CODE}) - 3]$ ,  $-3$  is the acceleration amplitude used to evaluate acceleration and deceleration behavior, because CODE 3 represents uniform driving behavior. This quantitative analysis method can consider the amplitude and change of acceleration at the same time, which can reflect the change of vehicle state and the change of vehicle control stability. Clearly, lower scores mean safer driving behavior.

$$\text{score}_{\text{comfort}} = \left[ \frac{1}{T} \int_0^T a_w^2(t) dt \right]^{1/2} \quad (3)$$

In the second part,  $\text{score}_{\text{comfort}}$  represents the root mean square value of weighted acceleration.  $T$  is the period (s) of vibration analysis;  $a_w(t)$  is the axial acceleration ( $\text{m/s}^2$ ); time history  $a(t)$  is the time history of weighted acceleration obtained by the weighted function filter  $w(f)$  of the corresponding frequency. This method refers to the evaluation standard of vehicle comfort formulated by ISO and only considers the influence of longitudinal acceleration on vehicle comfort. Lower scores mean a more comfortable driving.

$$\text{score}_{\text{economy}} = \begin{cases} \text{MOE}_e = e^{\sum_{i=0}^3 \sum_{j=0}^3 (L_{i,j}^e * v^i * a^j)}, a \geq 0 \\ \text{MOE}_e = e^{\sum_{i=0}^3 \sum_{j=0}^3 (M_{i,j}^e * v^i * a^j)}, a < 0 \end{cases} \quad (4)$$

In the third part,  $\text{score}_{\text{economy}}$  represents the fuel consumption of the vehicle. Here, VT-Micro fuel consumption model proposed by Virginia Tech University is adopted. This method can be used to calculate instantaneous fuel consumption based on vehicle speed and acceleration, and it is proved that the model results have a high degree of fit with the actual fuel consumption, where  $\text{MOE}_e$  is the instantaneous fuel consumption ( $L/s$ ) of the vehicle;  $v$  is the instantaneous speed of the vehicle ( $\text{km/h}$ );  $A$  is the instantaneous acceleration of the vehicle ( $\text{km/h/s}$ );  $L_{i,j}^e$  is the fitting coefficient of the instantaneous fuel consumption model when the vehicle is accelerating; the power of velocity  $v$  is  $i$  and the power of acceleration is  $j$ .  $M_{i,j}^e$  is the fitting coefficient of the instantaneous fuel consumption model when the vehicle slows down and the power of velocity  $v$  is  $i$  and the power of acceleration is  $j$ . The fitting coefficient can be obtained by querying the coefficient value table of the model. Lower scores for this segment mean more economical fuel consumption for vehicles. It should be noted that the overall driving behavior evaluation score consists of three evaluation index scores, and these evaluation indexes have different orders of magnitude and dimensions. If the original index data value is directly used for summation analysis, the results are easy to be unreliable; that is, the index with higher value will be highlighted, while the index with lower value level will be weakened. Therefore, the scores of different indicators need to be standardized.

#### 4. Driving Behavior Characteristics and Risk Analysis

**4.1. Acquisition of Experimental Data.** The data used in this paper was collected by the information acquisition equipment installed on the HDVs. The collection device includes a GPS module for collecting location data, an OBU module for

collecting vehicle movement data, and a cellular network module for uploading the collected data to a server. The collected information is uploaded to the Internet of vehicles service platform in real time. The parameters collected include the vehicle's GPS time, speed, direction, mileage, position, longitude, latitude, ACC state, and positioning state, etc. The data sampling frequency is 1 Hz.

The natural driving data of 10 HDVs were collected, including driving behavior data. The driving environment of these experimental vehicles includes different grades of roads such as highway, national highway, provincial highway, county highway, and rural highway. The professional drivers are generally older and the average daily driving distance of vehicles is more than 100 kilometers. There is a one-to-one correspondence between the driver and the vehicle. Data collection period is from June 29, 2019, to July 29, 2019, a total of 30 days; more than 6 million pieces of driving data are collected. The drivers in the experiment did not have fixed working hours or rest days, and most of them drove during the day. The working period was from 7:00 AM to 17:00 PM. The data collection time was long enough and of high quality to reflect the driving behavior of individual drivers and meet the needs of the experiment.

For the practicality of the proposed approach in real life, meanwhile, considering the positioning data is the most widely used and has high precision with large-scale application in real traffic until now; in addition, the acceleration is a typical index closely related to driving safety; the change of acceleration also represents the change of speed, which is another important index of driving safety. Therefore, in the analysis of driving behavior, this study mainly uses the longitudinal acceleration data of drivers.

**4.2. Driving Safety Preassessment.** Before analyzing the driving behavior pattern, this study needs to make a prejudgment on the driving situation of individual drivers, that is, to preliminarily evaluate driving safety by identifying the frequency of risk driving events. In this paper, two risk driving behaviors of sharp acceleration and sharp deceleration are considered, and the acceleration threshold value of  $\pm 0.3g$  proposed by Fazeen in this paper is used to detect these two driving behaviors. Using these thresholds, a quantified scoring method was used to assign each driver a driving safety label. Firstly, the proportion (percentage) of individual drivers' accelerated driving behavior (acceleration greater than  $3\text{ m/s}^2$ ) and decelerated driving behavior (acceleration less than  $-3\text{ m/s}^2$ ) in the total driving behavior was calculated. The total proportion of drivers is high, indicating that the driver operates frequently with rapid acceleration and has relatively low safety in driving behavior. The evaluation results are shown in Figure 2.

As can be seen from Figure 2, the accelerated or decelerated behaviors of different drivers account for a low proportion of all driving behaviors. The proportion of rapid deceleration behavior of most drivers is higher than that of rapid acceleration behavior. The proportion of acceleration behavior and deceleration behavior of different drivers conforms to the same trend. The driver with a high proportion of risk driving behavior has a tendency of more risk

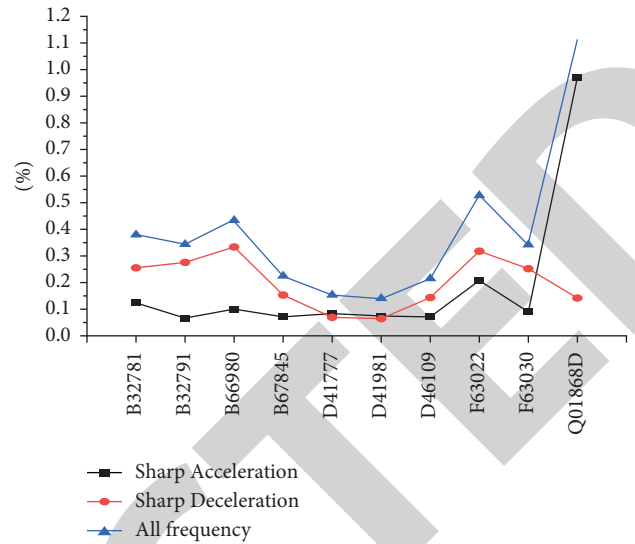


FIGURE 2: Overall proportion of drivers' risk driving behavior.

driving behavior. For example, the driver Q01868D has a higher proportion of accelerated behavior than other drivers. However, drivers with a low proportion of risk driving behaviors can temporarily consider their driving behaviors to be relatively safe, such as D41777 and D41981. In order to facilitate the subsequent comparative analysis, we divided the ten drivers into two groups. Assume that the top 5 with the highest proportion of total behaviors of rapid acceleration and deceleration include B32781, B32791, B66980, F63022, and Q01868D as drivers in the relatively dangerous group. The 5 drivers with a low proportion of rapid acceleration and deceleration behavior, including B67845, D41777, D41981, D46109, and F63030, were taken as the drivers in the relative safety group. The following will explore the personalized driving styles of different drivers more specifically by analyzing the driving behavior pattern.

**4.3. Integrated Driving Behaviors.** As can be seen from Figures 3 and 4, the most frequent driving behavior pattern of all drivers was cruising "333." The driving behavior pattern of individual drivers can clearly understand the driving styles of different drivers and their behavioral preferences in some driving situations. The shape and distribution of driving patterns can reflect the driving behavior characteristics of drivers, such as D41981 and Q01868D. The degree of dispersion and fluctuation of the driving behavior pattern of the two drivers can obviously reflect the degree of driving aggressiveness of the drivers. Some driving preferences of drivers can also be reflected in the figure. For example, compared with other drivers, driver Q01868D have high ranking of driving patterns "222" and "444." The ranking of risk driving behavior patterns such as rapid acceleration and deceleration and the amplitude of acceleration and deceleration fluctuation can reflect the driver's driving behavior and potential danger, but some drivers with a more stable overall style also have risk driving behavior patterns. That is, drivers in the relative safety group had more risk driving behavior patterns (e.g., "000") than those



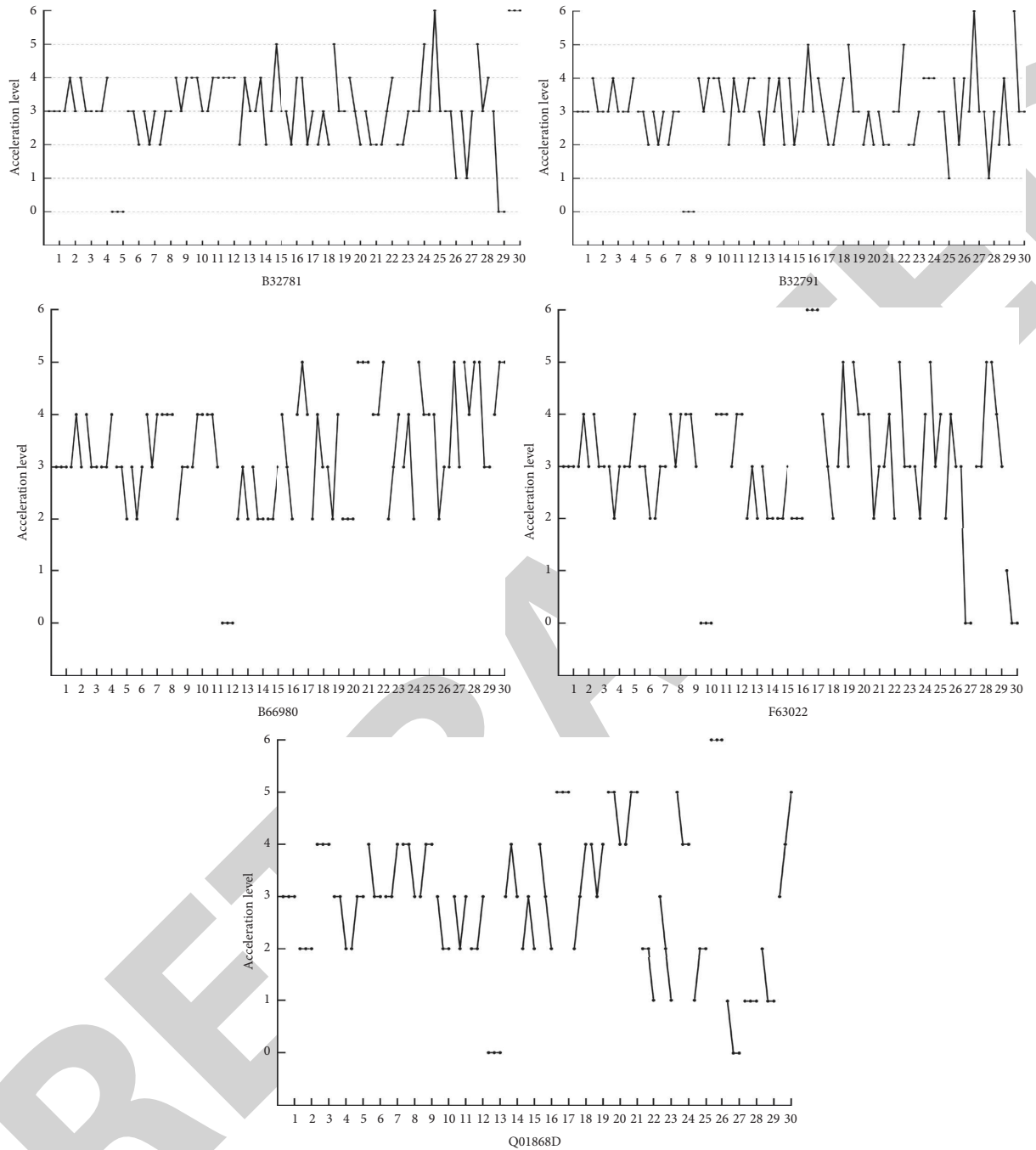


FIGURE 3: Dangerous group driving behavior pattern diagram.

in the relative safety group. In addition, the low proportion of rapid acceleration and deceleration in the overall driving behavior does not mean that the driver’s driving style is stable. It is necessary to judge the driver’s driving style by combining the type and distribution of specific driving behavior patterns. It also illustrates the importance of judging individual drivers’ driving styles.

4.4. Accelerated Driving Behavior. As can be seen from Figures 5 and 6, each chart represents the ranking from high to low of the occurrence frequency of the

corresponding driver’s accelerated driving pattern. The acceleration pattern diagram represents the similarities and differences and preferences of each driver in the acceleration pattern he or she chooses when taking the accelerated driving behavior. As can be seen from the figures, the occurrence frequency and ranking of the fast accelerating driving patterns defined as dangerous drivers, such as “555,” “666,” are generally higher than those of safe drivers. However, some safe drivers ranked higher than dangerous drivers in the accelerated driving pattern (e.g., mode “666”). For another example, for driver F63030, his rapid

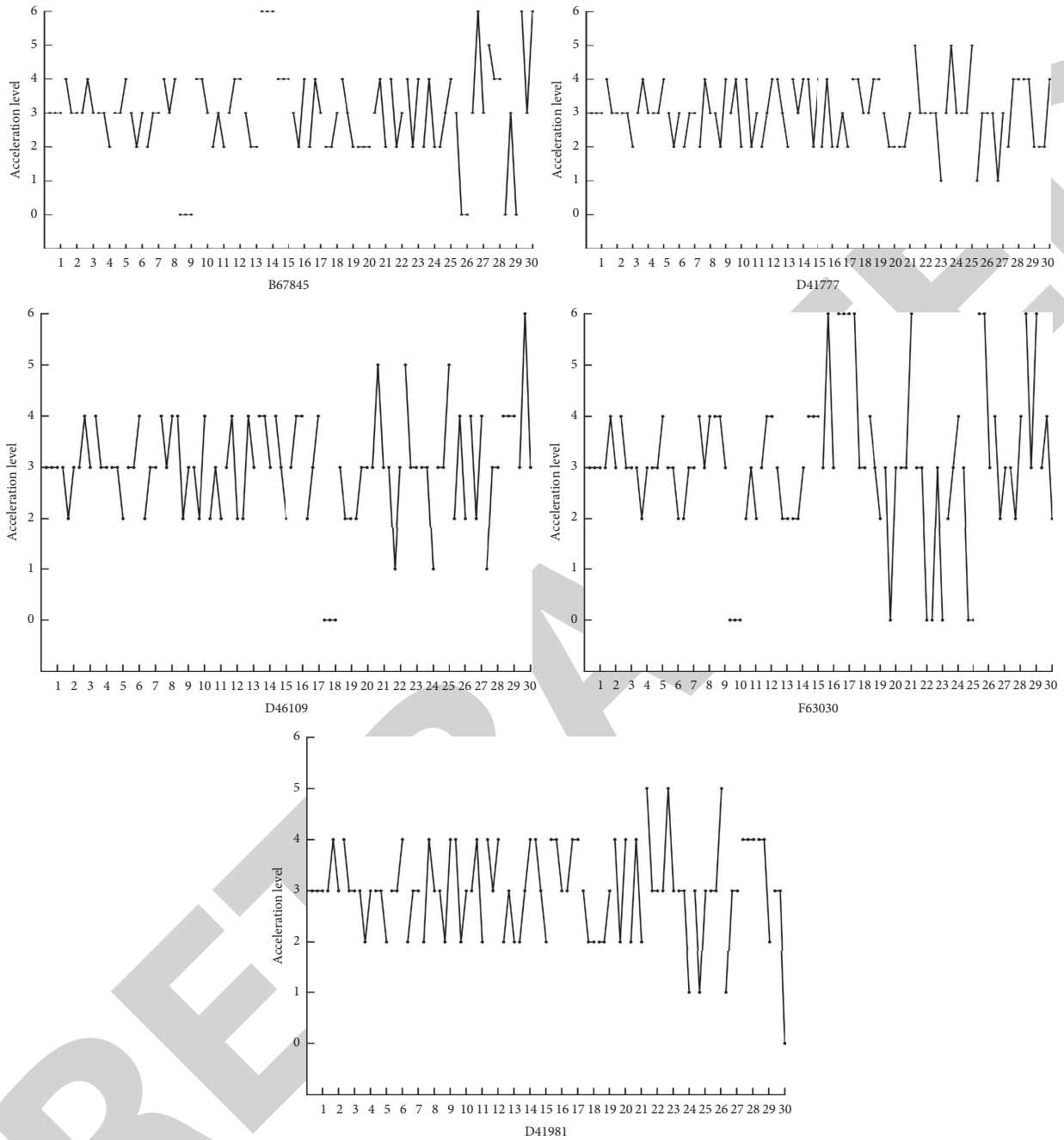


FIGURE 4: Safety group driving behavior pattern diagram.

acceleration and deceleration frequency is not high in the proportion of the overall driving behavior, but the driver's preference in acceleration is relatively aggressive. This shows that the overall proportion of risk driving behavior can only reflect the driver's driving risk on one hand. The reason is that the increase of the frequency of safe driving mode will reduce the overall frequency proportion of some risk driving modes, leading to the concealment of some risk driving behaviors. Some drivers generally have a smooth driving style but are prone to violent driving behavior at certain moments. Although the frequency of these risk driving behaviors is not high, the problem of increasing

accident probability caused by these behaviors cannot be ignored, which reflects the tendency of risk driving behaviors of drivers. Rankings and frequencies can complement each other to accurately differentiate the driving styles of different drivers.

It can be seen from the acceleration patterns of several drivers who are defined as safe drivers that with the increase of safety degree (the decrease of the overall proportion of drivers' acceleration and deceleration behaviors), such as "555," "666," the occurrence frequency and ranking of this dangerous pattern significantly decreased, and the acceleration behavior became more and

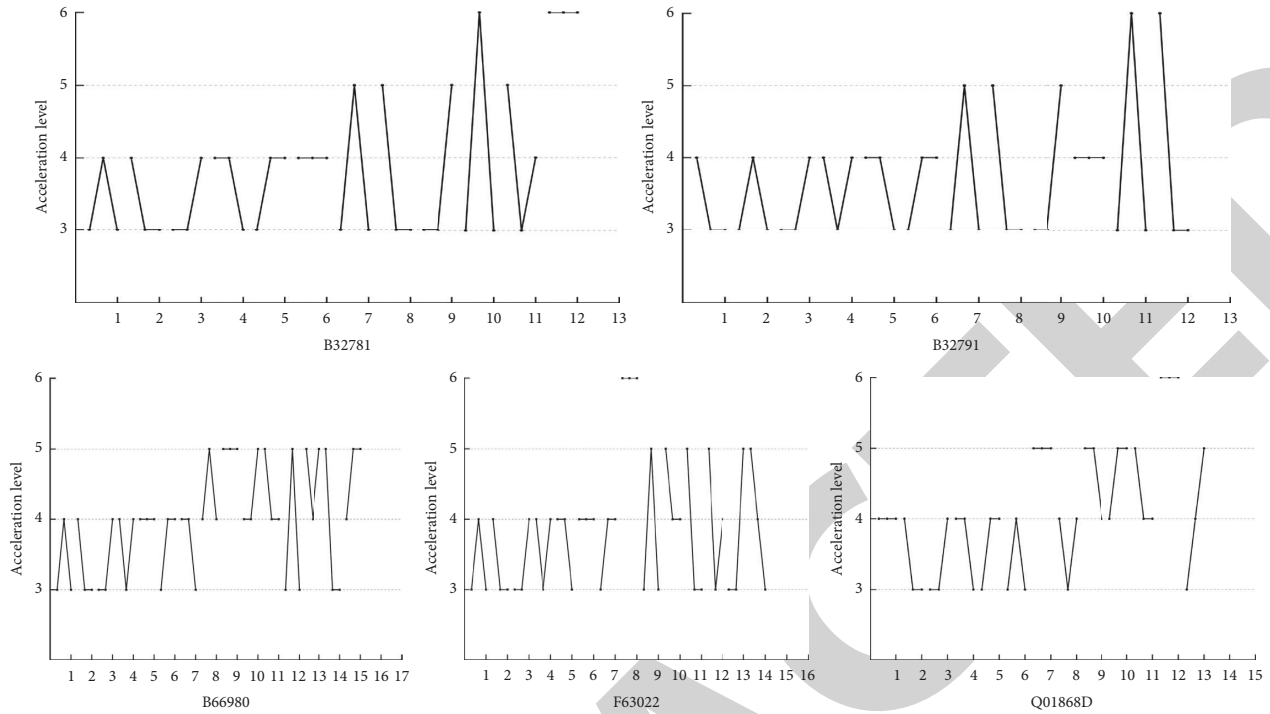


FIGURE 5: Dangerous group driving behavior pattern diagram.

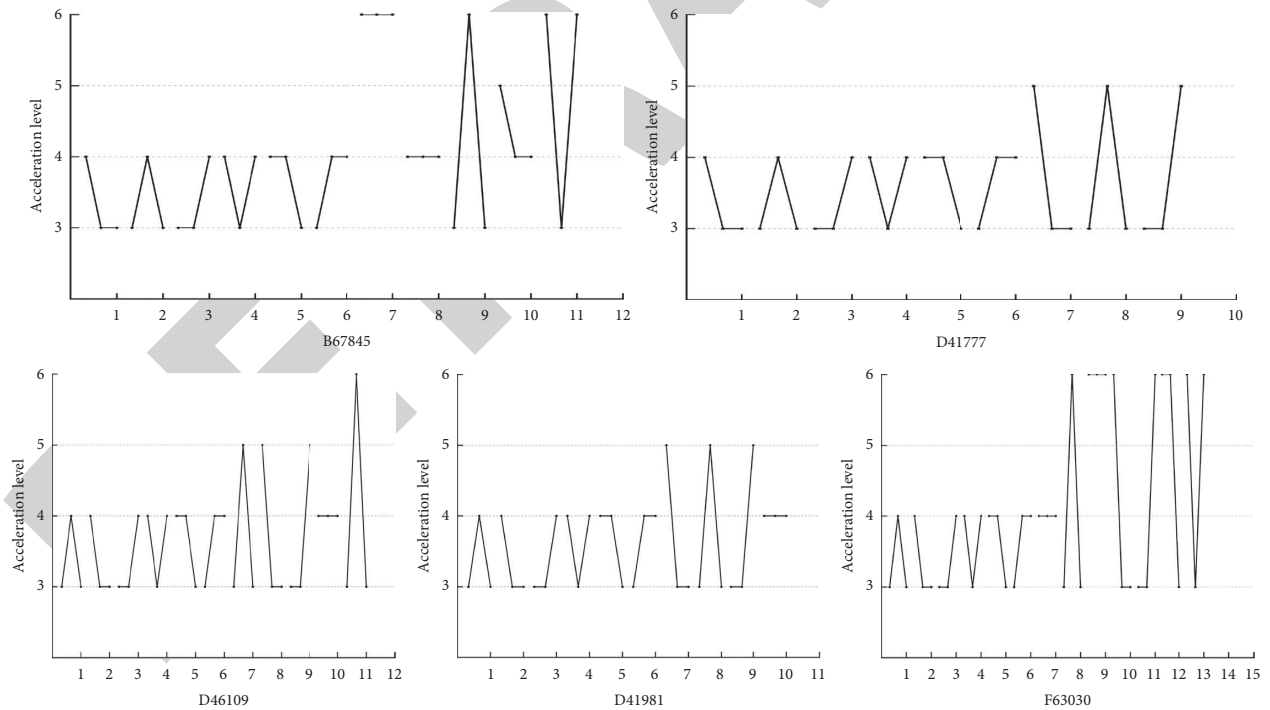


FIGURE 6: Safety group driving behavior pattern diagram.

more stable. In fact, in addition to the actual driving pattern occurrence frequency, the interaction between the “555” and “666” patterns and other patterns may also affect the driver’s driving safety level. In addition, the degree of driving safety is determined by acceleration and deceleration behavior, not just acceleration.

4.5. *Decelerated Driving Behavior.* As can be seen from Figures 7 and 8 for the occurrence frequency of the driver’s deceleration driving behavior pattern, the deceleration driving behavior pattern diagram represents the similarities and differences and preferences of each driver in the deceleration driving behavior. The pattern “000” and

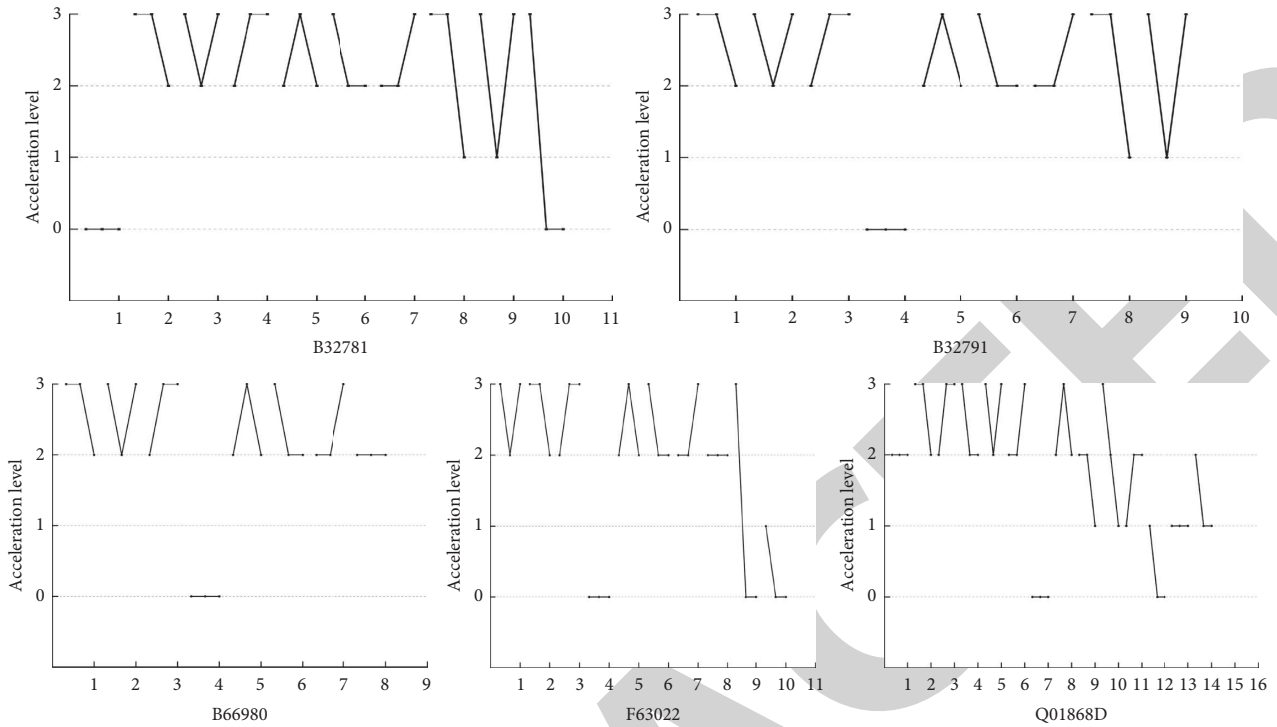


FIGURE 7: Dangerous group driving behavior pattern diagram.

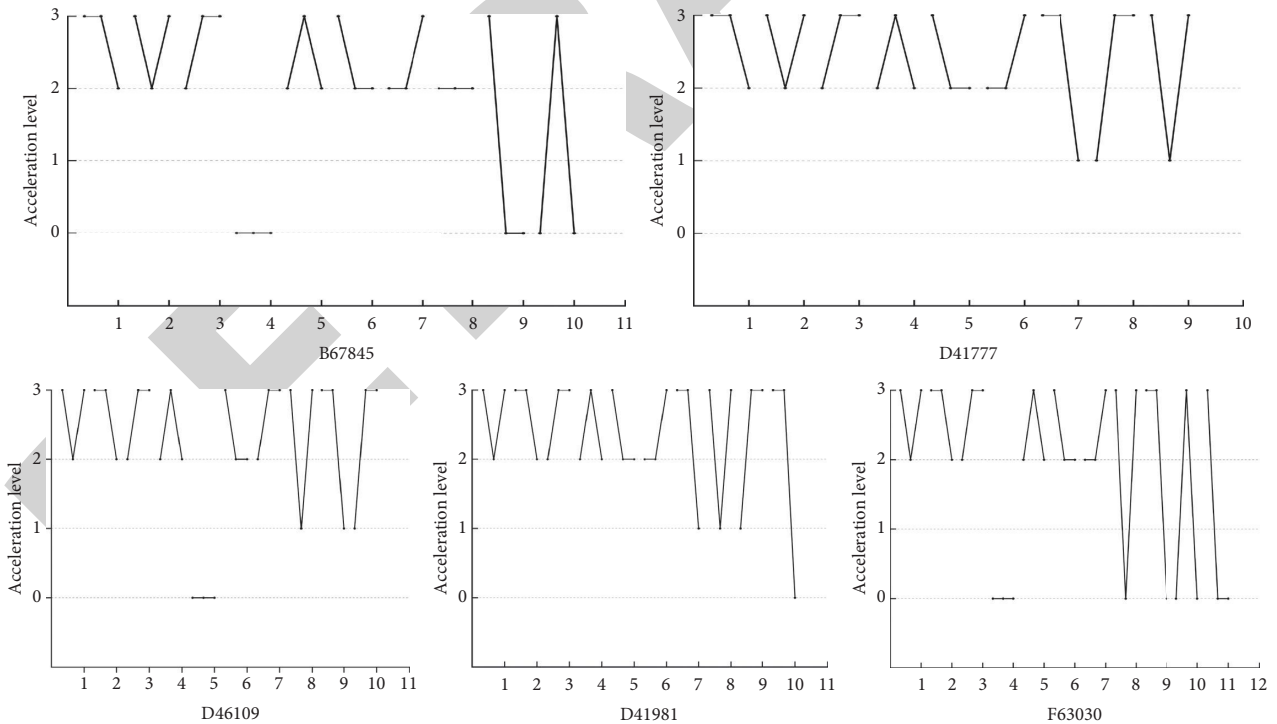


FIGURE 8: Safety group driving behavior pattern diagram.

deceleration patterns with large acceleration variations are good indicators of the driver’s style in deceleration driving behavior. The presumed danger group ranked higher in the “000” patterns, and some of their driving patterns reflected how hard they drove when slowing down. Drivers assumed

to be in the safety group also experienced sharp deceleration patterns, such as mode “000,” but were generally more stable. It should be noted that the deceleration driving behavior pattern of B67845 and F63030 in the safety group indicates that their deceleration driving behavior is risky.

The reason is that the driving pattern “000” ranks high, and there is a more aggressive driving pattern with large acceleration fluctuation. The F63030, in particular, was even more aggressive than some dangerous drivers in slowing down. In the previous two parts, it is also mentioned that the driver’s integrated acceleration and deceleration behavior is not much, but in the face of acceleration and deceleration driving behavior, the driver’s aggressive style, there is a potential accident risk. According to the deceleration pattern of different drivers, the driving preference and possible risk of each driver can be judged better.

**4.6. Driver Acceleration Level Distribution.** In order to further explore the acceleration preference adopted by drivers in acceleration and deceleration behavior, this section excludes drivers’ uniform driving behavior (acceleration level 3) and calculates the proportion of acceleration and deceleration behavior of each level in the integrated acceleration and deceleration behavior. Based on the radar chart, the behavior data of different drivers are visually evaluated, and the results are shown in Figures 9 and 10.

The hexagons of the radar map represent different levels of acceleration. Level 3 is removed here to exclude the influence of a large number of uniform driving behaviors on the acceleration and deceleration behavior proportion. The percentages of each driver’s different levels of acceleration are shown on the radar chart. The radar map clearly shows drivers’ preferences at various levels of acceleration. It can be found from the radar maps of the two groups that the occurrence ratio of dangerous group drivers is generally higher than that of safety group drivers for the more severe acceleration behavior (acceleration level 6) and the more severe deceleration behavior (acceleration level 0). For acceleration level 5 and acceleration level 1, the average proportion of drivers in the danger group was higher than 10% and close to 15%. The safety group averaged closer to 10 percent. The distribution in the figure shows that the acceleration and deceleration behaviors of drivers in the safety group are generally more gradual than those in the dangerous group. However, there are certain exceptions in each group. The acceleration and deceleration behaviors of drivers F63030 in the safety group were extreme, with more slight acceleration and deceleration behaviors (grade 4 and grade 2) and severe acceleration and deceleration behaviors (grade 6 and grade 0), which were reflected in the figure as hourglass. The acceleration and deceleration behaviors of drivers in the dangerous group Q01868D also tended to moderate, especially the slight deceleration.

**4.7. The Influence of Road Administrative Grade on Driving Behavior Pattern.** Road administrative grade has a certain influence on the driving speed and acceleration and deceleration behavior of drivers. In order to further explore the influence of road administrative grade on drivers’ driving behavior, this study selected a driver who had roughly the same driving time in the three road grades of national highway and above, provincial highway, prefectural highway, and below as a case to analyze. The driving behavior

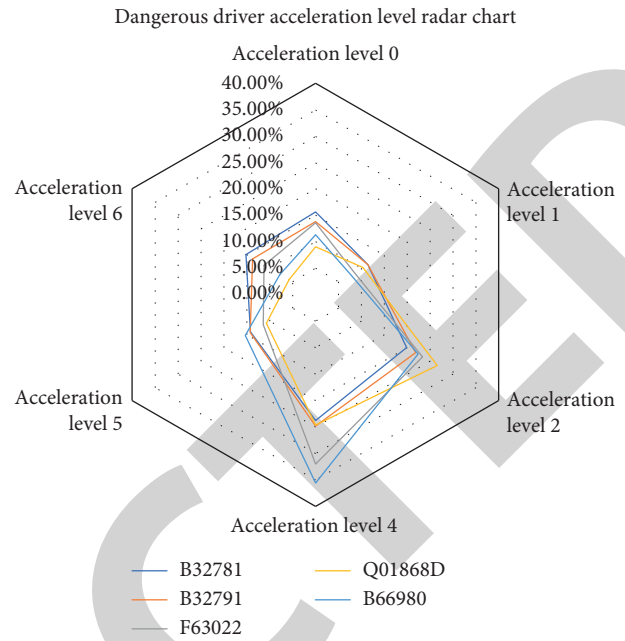


FIGURE 9: Acceleration rating of dangerous group drivers.

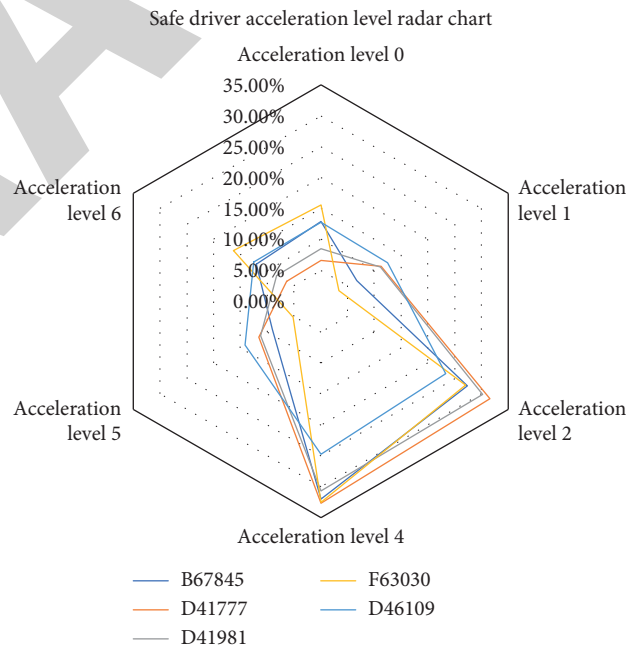


FIGURE 10: Acceleration rating of safe group drivers.

patterns of the driver under three different road administrative levels were compared, as shown Figure 11.

With the improvement of road administrative grade, the driver’s driving behavior becomes more stable; that is, the relatively safe driving mode ranks higher and occurs more frequently. The driving behavior pattern with higher potential risk occurred more frequently and ranked lower. For example, in driving mode “000,” it can be clearly seen that as the administrative level of the road is lowered, the driving mode’s ranking keeps rising. Driving pattern “000” is a typical driving behavior pattern with certain characteristics, which reflects the

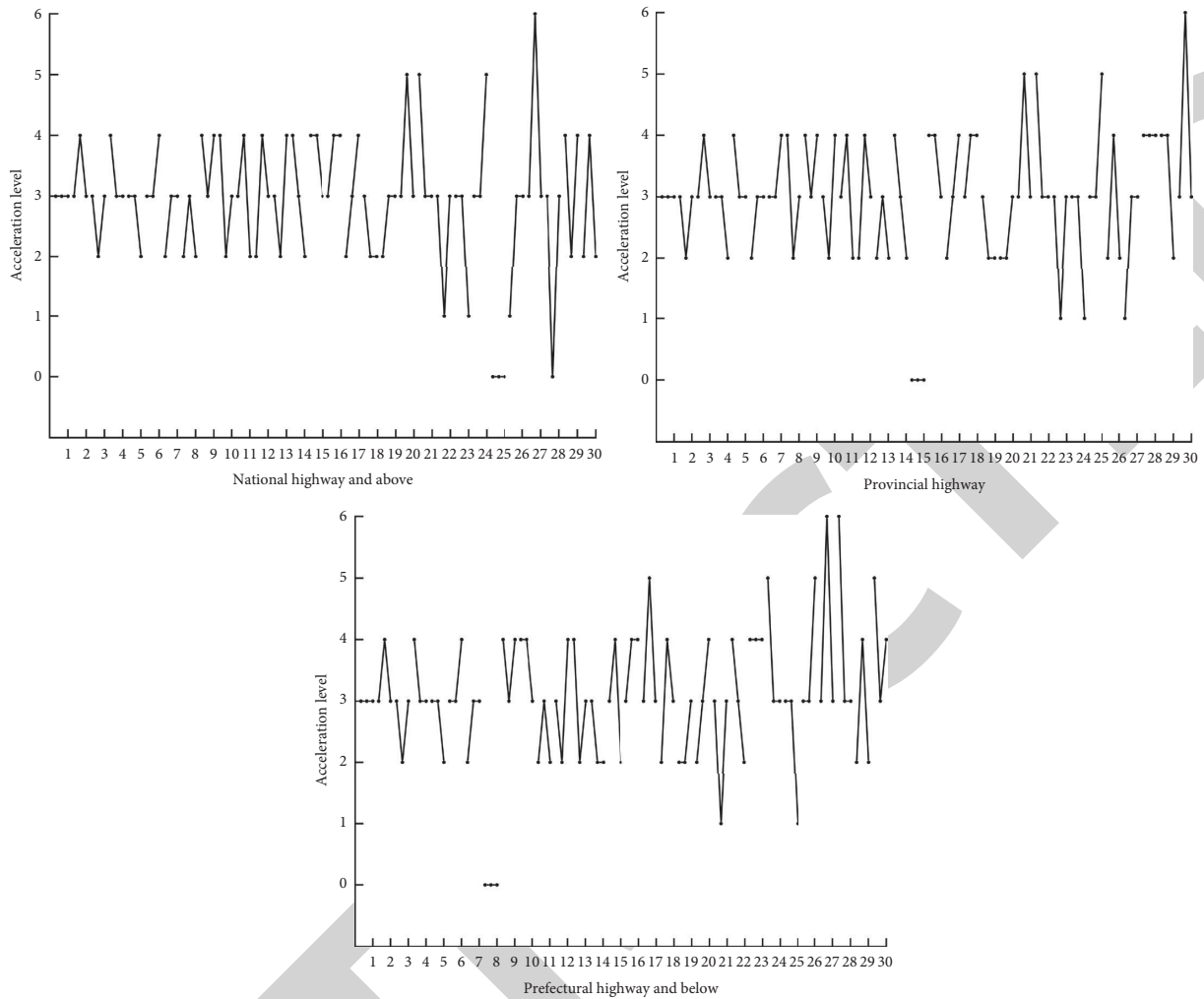


FIGURE 11: Patterns of driving behavior under three administrative levels of roads.

driver's driving preferences and habits. Under the conditions of three different administrative levels of roads, the driving patterns (e.g., driving pattern "000") with individual characteristics of drivers all appear, but the ranking of the modes in their respective road conditions is different. This also reflects that the driver's driving style is relatively fixed and he can maintain his own driving style in different road conditions. The driving behavior pattern diagram can not only show the distribution of driving behavior, but also clearly present the driving styles and driving risks of drivers under different road grades. Compared with the traditional method of fixed threshold, the driving behavior pattern diagram can avoid the problem that the fixed threshold is not suitable for drivers in different situations and can show the driver's unique driving habits without being influenced by the environment.

**4.8. Quantitative Assessment of Driving Risk.** The driving behavior of drivers was comprehensively evaluated by the comparative evaluation model of driving risk. This section divides the driver's driving behavior into three parts: overall driving behavior score, accelerated driving behavior score, and decelerated driving behavior score.

As can be seen from Figure 12, the scores of different behaviors of different drivers reflect the level of drivers. The higher the score, the higher the risk of driving behavior, and the lower the comfort and economy of driving. In general, the scoring trend of different drivers is roughly the same as the ranking trend of drivers' behavior scale graph based on fixed threshold. Drivers with a high proportion of risk driving behavior scored higher on acceleration, deceleration, and overall driving behavior. Those who drove more consistently, with less risky driving, tended to have lower scores. But there are some exceptions. The two drivers, B32781 and B32791, had a high proportion of overall risk driving behaviors, but their overall evaluation score was low. According to the data, this is because the driving behavior patterns of these two drivers are scattered, and the proportion of the first 30 typical driving patterns in the overall driving pattern is lower than that of other drivers, resulting in a lower score. The reason why the driving behavior evaluation score chart and the driving behavior ratio chart based on the threshold are not completely consistent is that there is still a lack of a clear definition of potential risks in the safety analysis of driving behavior. This method can describe the driver's personalized driving behavior characteristics,



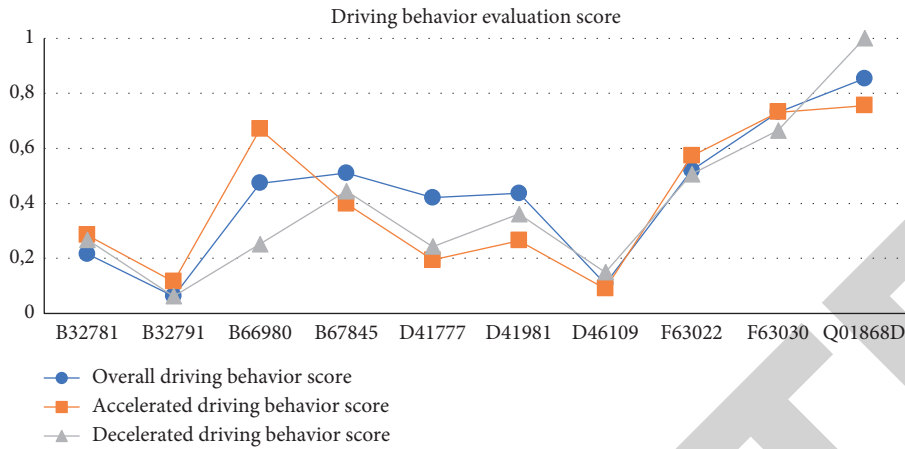


FIGURE 12: Driving behavior evaluation score.

comprehensively consider three factors of vehicle safety, driving comfort, and fuel economy, and quantitatively evaluate and compare the driver's driving risk.

In addition, the overall driving behavior score was different from that of only acceleration or deceleration. This is due to the difference in the pattern structure and frequency, and some drivers have more acceleration behavior or deceleration behavior. This phenomenon is acceptable. As can be seen from the figure, there are similar trends in the scores of these three behaviors among different drivers.

## 5. Discussion and Conclusion

According to the positioning data in the Internet of HDVs platform environment, this paper proposes an approach to evaluate the driver's driving risk and driving style with undefined threshold. Based on the vehicle driving data and the time series change of driver's driving behavior, the driving behavior pattern diagram was established. Through the driving behavior pattern diagram and driving risk modeling evaluation, the personalized driving behavior risk of different drivers was quantitatively evaluated, and the driving styles of different drivers were qualitatively and quantitatively compared. The main work and conclusions of this paper are as follows:

- (1) A graphic qualitative and quantitative analysis of individual driving behavior is carried out based on the change of driving behavior time series. Based on the positioning data of the actual HDVs, the driving behavior pattern was established to evaluate the driving behavior style and safety degree. The method is to evaluate the undefined threshold of the driver's driving style. Based on the vehicle driving data, a driving behavior pattern was established by using the driver's driving behavior changing with time.
- (2) Based on the driving behavior pattern diagram, the driving style of drivers is analyzed and evaluated qualitatively and quantitatively. The evaluation results show that the driving behavior pattern can reflect the driver's driving habits and potential risks. The ranking of risk driving behavior patterns such as rapid

acceleration and deceleration and the amplitude of acceleration and deceleration fluctuation can reflect drivers' driving habits and risks. Drivers with a more stable overall style also have some risk driving patterns. It is necessary to judge the driver's driving style by combining the type and distribution of specific driving modes. Even less risk driving behaviors may lead to the occurrence of driving accidents, and these risk driving behaviors are easy to be hidden by the overall driving data. Drivers will have different preferences in the degree of acceleration and deceleration of their driving behavior when faced with acceleration or deceleration. The driver's acceleration grade distribution also further reflects the driver's driving habit and style. In general, drivers with a low proportion of rapid acceleration and deceleration will adopt more moderate driving behavior and lower acceleration level in the face of acceleration and deceleration. The driving behavior pattern diagram can reflect the distribution of driving behavior and clearly present the driving style, habits, and driving risks of drivers.

- (3) According to the frequency and ranking of typical driving modes of drivers, a risk assessment modeling method for driving behavior was proposed from three aspects of vehicle safety, driving comfort, and fuel economy, and the similarities and differences of driving behaviors of different drivers were quantitatively analyzed. The evaluation results show that drivers with a high proportion of risk driving behavior scored higher on acceleration, deceleration, and overall driving behavior. Those who drove more consistently, with less risky driving, tended to have lower scores. But there are some exceptions.

## Data Availability

The data used in this paper were collected by the information acquisition equipment installed on the HDVs. The collection device includes a GPS module for collecting location data, an OBU module for collecting vehicle movement data, and a cellular network.

## Retraction

# Retracted: A DSGE Decision Model for Investigating the LPR Transmission Effect

### Computational Intelligence and Neuroscience

Received 17 October 2023; Accepted 17 October 2023; Published 18 October 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] X. Yang, "A DSGE Decision Model for Investigating the LPR Transmission Effect," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 2981558, 10 pages, 2022.

## Research Article

# A DSGE Decision Model for Investigating the LPR Transmission Effect

**Xing Yang** 

*School of Economics and Management, Beijing Jiaotong University, Beijing 100044, China*

Correspondence should be addressed to Xing Yang; 14113093@bjtu.edu.cn

Received 26 October 2021; Revised 3 January 2022; Accepted 4 January 2022; Published 21 January 2022

Academic Editor: Mario Versaci

Copyright © 2022 Xing Yang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In August 2019, People's Bank of China launched the reform of Loan Prime Rate (LPR) quotation formation mechanism and then made continuous progress in the order of “new loans first, followed by exiting loans,” dredging the interest rate transmission channel of “policy interest rate, LPR, loan interest rate.” In 2020, Chinese financial institutions have mainly referred to LPR pricing for loans, and the marketization level of loan pricing has significantly improved. This paper analyzed the policy effects transmitted by LPR through constructing a Dynamic Stochastic General Equilibrium (DSGE) decision model, and it was found that the financial market structure, pricing ability of commercial banks, and the degree of LPR application all affected the policy rate transmission effect and had an impulse impact on macroeconomic growth. Based on the above analysis, this paper proposed policy suggestions on the path of interest rate market-oriented reform and coping measures of commercial banks in China.

## 1. Introduction

In August 2019, the market-oriented reform of China's interest rate was further deepened. People's Bank of China launched the reform of the formation mechanism of LPR in the loan market and continued to promote it in the order of “new loans first, followed by exiting loans,” forming a rate transmission mechanism of “policy interest rate, LPR, loan interest rate.” By the end of 2020, Chinese financial institutions have mainly referred to LPR pricing for loans, and the marketization level of loan pricing has significantly improved. However, it is worth noting that China's financial system structure is dominated by indirect financing, and there are some problems such as information asymmetry, implicit high leverage, and inadequate expectation management in the financial market. It is difficult to give full play to the profit-seeking function of capital, and the actual impact of interest rate reform on economic output is worth discussing. From the perspective of international experience, through the reform of interest rate marketization, it is conducive to narrowing the interest rate corridor, reducing short-term interest rate fluctuations, enhancing the sensitivity of real economic deposits interest rate and loan interest

rate to policy interest rate, improving the transmission efficiency of monetary policy between different interest rates, and strengthening financial services in the real economy. Under the background of the imperfect system of China's financial market, after the reform of the LPR pricing mechanism, how efficient is the transmission efficiency of monetary policy? What is the actual impact on economic output? These are all issues that need to be studied. The study of the above problems will help us to further dredge the transmission path of monetary policy in a targeted manner, reduce the financing cost of the real economy, force the market-oriented development of the pricing method of funds of commercial banks, and further improve the efficiency of macroeconomic regulation and control.

Therefore, it is of great significance to study the transmission effect of LPR pricing mechanism and its impact on economic growth. At present, the research mainly focuses on the effect evaluation of quantitative control tools and price control tools of monetary policy and the effect evaluation of interest rate regulation and interest rate market-oriented reform, and some studies believe that too fast interest rate market-oriented reform will increase the risk of financial institutions. However, there is a lack of research on the

transmission effectiveness of LPR interest rate reform and the impulse impact on macroeconomic growth under different degrees of marketization. In order to further explore these questions, the paper constructs a DSGE model with financial market friction, studies the marketization degree in the process of LPR reform by difference scenarios, and analyzes the effectiveness of models such as calibration analysis, parameter estimation, and impulse response, so as to provide objective theoretical support for a comprehensive understanding of the policy effect of LPR.

## 2. Literature Review

The study on the transmission effect of monetary policy rate began as early as the late 19th century. Tobin [1] and Modigliani [2] pointed out that monetary policy transmission relied on not only interest rate transmission channels, but also nonmonetary assets price channels, such as enterprise Tobin's Q channel (assets price), the resident's wealth effect channel, and the exchange rate channel for transmission. Relying on the IS-LM model, Bernanke and Blinder [3] and Bernanke and Gertler [4] explained that not only the total money supply from the central bank, but also changes of bank credit supply quantity, both have a significant impact on monetary policy transmission. Due to incomplete information and agency costs in financial markets, expansion or shrinkage of bank credit quantity can cause economic agents to make different financing decisions and affect economic behaviors such as investment, consumption, etc. According to the above viewpoints, Mishkin [5] divide the monetary policy transmission channels into the interest channel, the credit channel, the assets price channel, and the exchange rate channel. Since then, a large number of empirical studies have begun to quantitatively measure the transmission efficiency of different channels. For example, Evans and Mashall [6] used Vector Autoregressive model (VAR) to study the role of monetary policy and inflation expectation on the term structure of interest rate and pointed out that tight monetary policy had a positive but temporary effect on short-term interest rate, while long-term interest rate was almost not affected. De Bondt [7] studied the European banking sector and found out through analysis that the market structure of the banking sector affected the transmission effect of monetary policy to some extent. The highly monopolistic banking would weaken the transmission effect of monetary policy, while market-oriented banking would make commercial banks more sensitive to interest rate changes and improve the transmission effect of the central bank monetary policy. Bolton and Freixas [8] further indicated that the capital adequacy ratio rules of commercial banks, while enhancing the risk prevention ability of commercial banks, reduced the ability of commercial banks to derivate money and weakened the transmission effect of the central bank's monetary policy in the commercial bank system. Milani and Treadwell [9] and Gomes et al. [10] show that the monetary policy with full expectation management has greater degree of macro-control on the economy and longer attenuation period of impact effect. Empirical studies on DSGE models based on

open economy done by Meinus and Tillmann [11] and Ge [12] manifest that the transmission efficiency of price monetary policy shows a significant improvement trend during the sample period, while the macrocontrol effect of quantitative monetary policy is gradually weakening.

At present, the domestic literature on interest rate transmission mainly focuses on the transmission effect study of traditional loan and deposit benchmark rates, while there are few studies on transmission effect of LPR reform. Yingkun Jiang et al. [13] argued that the bank credit transmission channel played an important role in China's monetary policy transmission channels. However, the constraints of China's traditional credit system, as well as the state-owned nature and limited competition of commercial banks, can slow down the central bank's interest rate policy transmission and cause efficiency losses. Dong He and Honglin Wang [14] pointed out that, different from developed countries, China gradually formed a monetary policy framework with the dual interest rate system as the main feature in the market-oriented reform. China transmitted the monetary policy to market interest rate through various policy instruments and realized the monetary policy regulation with the credit quantity. Li Ma et al. [15] showed that the deposit reserve ratios of commercial banks were important factors that restricted the interest rate policy transmission effect of commercial banks. Deposit reserve ratio adjustment could guide market changes and macro-economic development and promote the transmission effect of monetary policy by changing the adaptive expectations of market participants. Jun Ma et al. [16] constructed a DSGE decision model and pointed out that China's financial system, with indirect financing as the main body, faced many institutional constraints. The transmission of central bank's policy rate adjustment to each financial market rate would be influenced by China's financial market structure, loan scale restrictions, deposit reserve ratio, etc. Daoping Wang [17] used the panel data of 88 countries around the world to study the relationship between interest rate marketization and banking crisis which also shows that the improvement of marketization will significantly increase the probability of banking crisis. Ning Xu et al. [18] believed that China's LPR reform based on Medium-Term Lending Facility (MLF) tends to be more market-oriented and improves the transmission efficiency of monetary policy.

In summary, the existing research results showed that, due to China's monetary market structure with indirect financing as the main body, policy rate changes were mainly transmitted through bank credit. But the transmission mechanism was not smooth enough and the effectiveness of interest rate policy was low due to factors such as capital constraints of commercial banks, deposit reserve ratio constraints, and prudent management behaviors under imperfect competition. At present, the LPR interest rate has a typical market-oriented attribute, but its transmission efficiency and impact on economic growth are still lacking forward-looking empirical analysis. Based on the research methods of Zihan Huo et al. [19] and Long Zhang et al. [20] using DSGE model to analyze the impact effect of China's monetary policy on macroeconomy, this paper constructs a

DSGE model including five subjects: government, residents, enterprises, central bank, and commercial banks to perspective dynamic changes of the regulation effect of LPR channel, in order to provide reference for improving the regulation of monetary policy.

### 3. Research Mechanism

*3.1. Transmission Mechanism of Price-Based Regulatory Monetary Policy.* When a country's economy is in a rising phase, the actual GDP growth rate is often higher than the country's potential economic growth rate, which tends to cause a continuous rise in CPI and drives residents' inflation expectations, thus intensifying the inflationary trend. To prevent the overheated economic fluctuations, the country's government often adopts a tight monetary policy to hold down the inflationary trend. In the price-based monetary policy regulation, the monetary authority often stimulates the rise in the financial market interest rate through open market operations, which makes medium- and long-term funds flow into the capital market. The shortage of funds will prompt commercial banks to increase the loan rate and squeeze the momentum of enterprise investment, raising the enterprise investment cost and slowing down the trend of overheated economy.

When a country's economy is in a downward phase, the actual GDP growth rate is usually lower than the country's potential economic growth rate, resulting in a continuous downturn of the CPI and causing deflation. Therefore, it is required to implement an accommodative monetary policy to stimulate economic growth. In the price-based monetary policy regulation, the monetary authority often promotes the decline of financial market interest rate through open market operations to squeeze medium- and long-term capital outflows from the capital market. The sufficient funds will prompt commercial banks to reduce loan rate, which is conducive to enterprises to increase the investment scale, thus promoting economic recovery.

Since the mid-20th century, major developed countries have successively improved their monetary operation procedures by implementing a deposit and loan facility system, increasing the frequency of open market operations, narrowing the interest rate corridor range, as well as carrying out interest compensation on deposit reserves, adopting average reserve assessment, strengthening policy communication, and improving transparency, which can make money market interest rate deviate less from the central bank's target level and is more beneficial to the interest rate guidance of central bank. In October 2006, People's Bank of China introduced Shanghai overnight interbank offered rate (Shibor) by borrowing from London interbank offered rate (Libor), Euro interbank offered rate (Euribor), and other interest rates and made an active exploration in cultivating the market benchmark interest rate. However, due to the great fluctuation of Shibor in actual operation, combined with the single transaction structure of the money market, the borrowed funds basically showed a single flow direction (from large banks to small-sized and medium-sized banks), while large banks all belonged to the price quotation units of

Shibor interest rate, which was easy to form a monopolistic interest rate to some extent, and had low market recognition as a monetary policy benchmark interest rate. In this process, People's Bank of China continuously enriched the means of interest rate regulation, built long-term, medium-term, and short-term policy interest rate system including 3-month, 6-month, and 1-year MLF interest rate and improved the efficiency of monetary policy interest rate transmission.

*3.2. Four Stages of LPR Mechanism Transmission.* In October 2013, People's Bank of China launched the LPR, aiming to further stimulate market-oriented reform, improve the benchmark interest rate system, and guide the pricing of loan on credit market products. However, the LPR has been following the 1-year benchmark interest rate of loans announced by the central bank, and the interest rate level has been maintained at around 4.31%, which is much higher than the 1-year MLF interest rate of 3.30% and the 1-year Shibor interest rate of 3.52%. In August 2019, People's Bank of China further optimized the calculation method of LPR based on the mode of "new loans first, followed by exiting loans," adopted the way of 1-year LPR equal to 1-year MLF plus base point for loans, and expanded the range of quoting banks. From the current operation condition, the transmission can be divided into four stages: In the first stage, the central bank lowers the MLF quotation rate to influence the financial market interest rate. In the second stage, the quoting bank adopts the way of "MLF + base point" for the LPR quotation, which should be calculated and published by China Interbank Offer Center. In the third stage, commercial banks adjust the loan rate according to the LPR quotation. In the fourth stage, commercial bank lending rate affects the investment and financing decisions and costs of enterprises and residents, thus affecting macrotargets such as economic growth, inflation, and employment. The transmission path is MLF→LPR→commercial bank lending rate→real economy→economic target (Figure 1).

In 2019, only 20% of the loans of Chinese commercial banks were floating priced up and down according to the LPR quotation, and about 80% of the loans were still floating priced up and down according to the loan benchmark interest rate published by the central bank in the past. By the end of 2020, commercial bank loans had realized transformation from policy benchmark interest rate to market interest rate, so as to ensure that the policy operation intention of People's Bank of China can be effectively transmitted to other markets.

*3.3. Main Idea of DSGE Decision Model Construction.* The estimation of the DSGE decision model is generally divided into three steps: First, based on the economics theory and the economic operation characteristics in China, the optimal behavior characteristic of each economic agent when the economic system is in equilibrium is set. Second, the policy equation of the economic variables is solved, and the model parameters are estimated by running calibration and Bayesian estimation. Third, the impulse response is used to

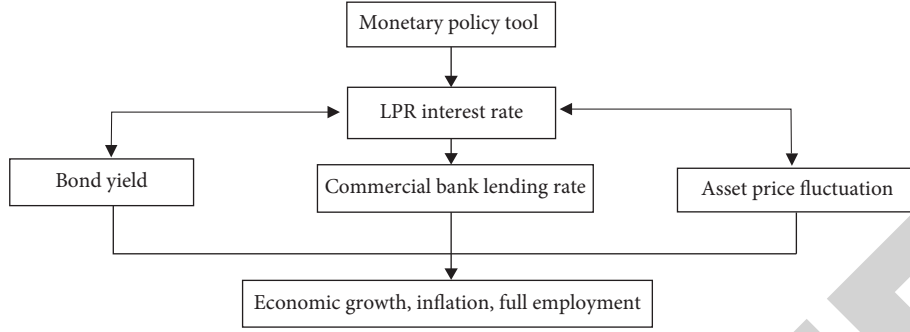


FIGURE 1: LPR mechanism transmission framework diagram.

carry out efficiency analysis of the model and test the explanatory ability of the model for the actual economic problems.

Foreign studies on the transmission mechanism of monetary policy mainly focus on how market rate changes affect economic operation through financial markets, while there are few studies on the impact of policy rate changes transmitted to the real economy. This is mainly because the financial systems of developed countries are often dominated by the direct financing market, where regulatory policies have fewer constraints on financial markets and benchmark market interest rate is mostly determined by transactions between financial markets. In contrast, China's financial system is dominated by the indirect financing market, where the policy rate changes have a significant impact on the real economy. Based on this, this paper constructs a DSGE model that includes five entities: government, residents, enterprises, central banks, and commercial banks, and defines the corporate loan interest rate as adjusting according to the LPR pricing level and according to the three scenarios of 50%, 80%, and 100% loans with reference to LPR pricing, to discuss the impact of the change of the unit policy interest rate on economic growth.

## 4. Research Method

### 4.1. DSGE Decision Model Frameworks of 5 Sectors

**4.1.1. Resident Sector.** The resident sector realizes the maximization of economic utility by supplying labor, earning income, and consumption. Its utility equation is as follows:

$$U\left(C_t, N_t, \frac{M_t}{P_t}\right) = Z_t^c \frac{C_t^{1-\sigma_c}}{1-\sigma_c} + Z_t^m \frac{(M_t/P_t)^{1-\sigma_m}}{1-\sigma_m} - Z_t^n \frac{N_t^{1+\sigma_n}}{1+\sigma_n}. \quad (1)$$

The expected utility is

$$\max: E_t \left[ \sum_{t=0}^{\infty} \beta^t U\left(C_t, N_t, \frac{M_t}{P_t}\right) \right]. \quad (2)$$

The constraint condition is

$$\text{s.t: } C_t + \frac{M_t}{P_t} = \alpha_t^s N_t^s + \frac{M_{t-1}}{P_t}, \quad (3)$$

where  $E_t$  is the mathematical expectation, and  $\beta^t$  is the discount factor ( $0 < \beta \leq 1$ ).  $C_t$  is the household consumption

level, and  $N_t$  is the labor supply of the household to society.  $M_t$  is the nominal assets balance held by the household, and  $P_t$  is the product price in the current period.  $M_{t-1}$  is the household's nominal assets balance held in period  $t-1$  and substituted into period  $t$ .  $(M_t/P_t)$  is the actual assets balance held by the household.  $Z_t^c$ ,  $Z_t^m$ , and  $Z_t^n$  are exogenous impact variables for consumption demand, labor supply, and holding assets, respectively.  $\sigma_c$ ,  $\sigma_n$ , and  $\sigma_m$  are the consumption demand elasticity, labor supply elasticity, and money demand elasticity, respectively.  $\alpha_t^s$  is the average wage level.

**4.1.2. Enterprise Sector.** The enterprise sector obtains the maximum enterprise welfare by investing capital and labor. The objective function is  $\max: E_t [\sum_{t=0}^{\infty} \beta_t^p \lg C_t^p]$ , where  $\beta_t^p$  is the discount rate of the manufacturer and  $C_t^p$  is the remuneration obtained by the business owner. Suppose the production equation of the manufacturer is

$$Y_t^p = A_t^p (K_t^p)^{\alpha_1} (N_t^p)^{1-\alpha_1}, \quad (4)$$

where  $Y_t^p$  is the output of the enterprise, and  $A_t^p$  is the production technology level of the enterprise.  $K_t^p$  is the capital invested by the enterprise, and  $N_t^p$  is the labor invested by the enterprise.  $\alpha_1$  is the output elasticity of capital input, and  $\alpha_1 \in (0, 1)$ .

Refer to the research of Xiangyang Li et al. [21]; this paper assumes that the resource constraint equations of the enterprise are as follows:

$$(1 - T_p)Y_t^p + L_t^p + (1 - \varepsilon)K_{t-1}^p - L_{t-1}^p \frac{1 + R_t^p}{\pi_t}, \quad (5)$$

$$C_r^p - K_t^p = 0,$$

$$L_t^p \frac{1 + R_t^p}{\pi_t} = m_p K_t^p,$$

where  $T_p$  is the tax rate to be paid by the enterprise, and  $\varepsilon$  is the depreciation rate.  $L_t^p$  is the loan balance obtained by the enterprise in period  $t$ , and  $R_t^p$  is the loan interest rate.  $m_p$  is the asset-liability ratio of the enterprise. Enterprise loans are generally restricted by the loan interest rate and the asset-liability ratio.



**4.1.3. Government Sector.** The basic behavior characteristics of the government are simply simulated in this model; that is, the government obtains funds through tax revenue and debts, and these funds are mainly used for government investment and government consumption, except for paying back debt and interest of the last period. The government budget constraint is

$$\frac{g_t}{p_t} + T_t + T_p Y^t = B_t + G_t, \quad (6)$$

where  $g_t$  is the funds received by the government from issuing treasury bonds, and  $P_t$  is the inflationary price of the current period.  $T_t$  is the tax revenue paid by the residents and  $T_p Y^t$  is the government tax revenue (where  $T_p$  is the tax rate and  $Y^t$  is the enterprise output).  $B_t$  is the government consumption expenditure, and  $G_t$  is the government investment expenditure.

**4.1.4. Central Bank.** People's Bank of China is now gradually abandoning the money supply as the intermediate target and more focusing on the adoption of policy rate. There are generally 4 ultimate objectives: price stability, promoting economic growth, maintaining employment, and balance of international payment. However, considering Mundell's Impossible Trinity, this paper mainly considered price stability, economic growth, and maintaining employment. Accordingly, the monetary policy constraint equation of People's Bank of China is as follows:

$$\frac{R_t}{R} = \left( \frac{R_t}{R_{t-1}} \right)^{\theta_R} \left[ \left( \frac{\pi_t}{\pi_{t-1}} \right)^{\varphi_\pi} \left( \frac{Y_t}{Y_{t-1}} \right)^{\varphi_y} \left( \frac{M_t}{M_{t-1}} \right)^{\varphi_m} \left( \frac{p_t}{p_{t-1}} \right)^{\varphi_p} \right]^{1-\theta_R} Z_t^r, \quad (7)$$

where  $(R_t/R_{t-1})$  is the changing value of policy rate (this paper adopted MLF interest rate).  $(M_t/M_{t-1})$  is the changing value of money supply and  $(\pi_t/\pi_{t-1})$  is the changing value of unemployment rate.  $(Y_t/Y_{t-1})$  is the changing value of total social output, and  $(p_t/p_{t-1})$  is the changing value of inflation rate. The parameters  $\theta_R$ ,  $\varphi_\pi$ ,  $\varphi_y$ ,  $\varphi_m$ , and  $\varphi_p$  are the elasticity of interest rate, unemployment rate, total output, money supply, and inflation from their steady-state deviation degree with respect to interest rate, respectively.  $Z_t^r$  is the exogenous impact variable of China's monetary policy.

**4.1.5. Commercial Bank.** Commercial banks mainly face Net Interest Margin (NIM) constraints in business operations. Commercial banks earn interest income from loans to cover the deposit interest cost and maintain interest spread income. Therefore, the budget constraint equation for commercial banks is as follows:

$$\begin{aligned} D_t r_t + F_t &= L_t^p R_t^p + B_t^G R_t^G, \\ R_t^p &= LPR_t + K_t, \end{aligned} \quad (8)$$

where  $r_t$  is the deposit interest rate, and  $D_t$  is the commercial bank deposit balance.  $F_t$  is the commercial bank profit, and  $R_t^p$  is the enterprise loan interest rate (LPR<sub>t</sub> plus floating

point  $K_t$ ).  $R_t^G$  is the government financing rate.  $L_t^p$  is the balance of enterprise loans and  $B_t^G$  is the balance of government financing.

**4.2. Data Selection, Parameter Calibration, and Economic Steady State.** We consider that  $Y = C_t + Y_t + F_t + G_t + R_t$ . Among them,  $Y$  represents the total output of the economy,  $C_t$  represents the resident sector,  $Y_t$  represents the corporate sector,  $B_t$  represents the commercial banking sector,  $G_t$  represents the government sector, and  $R_t$  represents the central bank sector. In order to further examine the effectiveness and efficiency of monetary policy transmission after the implementation of LPR reform, we will select the main observation data according to the DSGE model framework of the five departments and simulate the transmission of monetary policy by means of calibration analysis.

**4.2.1. Data Selection.** The sample interval for this section is from January 1980 to December 2021. The main observational variables include interest rate data (1-year MLF interest rate, 1-year LPR interest rate, 1-year bank deposit rate, 1-year bank loan interest rate, and 1-year treasury bond interest rate); bank data (bank deposit balance as a proportion of GDP, bank loan balance as a proportion of GDP, new loans to new social financing, new government bonds to new social financing, and bank net interest margin); economic data (GDP growth rate, inflation rate, unemployment rate, M2 growth rate, and corporate asset-liability ratio). The data are all from the WIND database. Here, considering the substitution of wealth management products for resident deposits, the deposit interest rate adopts the average expected yield of 1-year wealth management products of large commercial banks. This paper first deflates GDP, bank deposits, and bank loans, then uses the X12 method to seasonally adjust all data, and finally takes all the data logarithmic and first-order differences. The statistical description of the data is shown in Table 1.

**4.2.2. Parameter Calibration.** The steady-state value of the variable  $R_t$  is defined as  $R_t^*$ , and  $\widehat{R}_t$  is the steady-state deviation of the variable  $R_t$  with respect to  $R_t^*$ , that is,  $\widehat{R}_t = \log R_t - \log R_t^*$ . The linear policy equation of the DSGE decision model can be obtained by conducting first-order differential and first-order Taylor expansion. After that, the a priori calibration method and automatic model solution are required to obtain the estimated values of the parameters.

In this paper, Carlstrom et al. [22] study and calibrate some static parameters with relatively stable values and then focus on the Bayes method to estimate the dynamic structural parameters related to the degree of marketization and interest rate transmission. For example, this article uses  $\theta_R$  as an indicator of the degree of financialization; we estimated three marketization degree coefficients based on three samples of different periods, of which from January 1980 to December 2007, the marketization degree coefficient was 0.012, which was the interest rate control stage; from January 2008 to December 2018, the marketization degree

TABLE 1: Descriptive statistics of all variables.

Name	Obs	Mean	Std.	Min	Max
MLF interest rate	504	3.1341	0.5999	2.9500	3.3000
LPR interest rate	504	4.6842	1.7160	3.8000	5.7700
Deposit rate	504	4.9905	0.1464	3.000	30.0000
Loan interest rate	504	5.5948	0.8816	4.7500	7.5600
Treasury interest rates	504	2.6672	0.6976	0.8871	4.2503
Deposit balance/GDP	504	1.9433	0.2194	1.4102	2.3352
Loan balance/GDP	504	1.3778	0.2371	0.9579	1.9156
New loans/new social financing	504	0.6549	11.4001	0.4213	1.1563
New government bonds/new social financing	504	0.1143	6.7104	0.0171	0.2392
Bank NIM	504	2.3617	0.2622	2.0300	2.7700
GDP growth	504	0.1173	0.1026	0.0220	0.1423
CPI	504	2.6350	1.9774	-1.8000	8.7000
Jobless rate	504	4.0165	0.1647	3.6100	4.3000
M2 growth	504	13.0246	5.0742	8.0000	29.7400
Asset-liability ratio of industrial enterprises	504	0.6103	2.0045	0.5830	0.6420

coefficient was 0.4523, and the market benchmark interest rate system represented by Shibor was constructed, which was the market cultivation stage; from January 2019 to December 2021, with the advancement of the reform of China's LPR interest rate pricing system, the degree of marketization coefficient was 0.6851, which was the stage of gradual market liberalization. Considering the above factors, the final calibration value of  $\theta_R$  is 0.4832. The estimation results are shown in Table 2.

**4.2.3. Steady-State Simulation.** This section uses the impulse response analysis of the DSGE model to explore the role of changes in the benchmark interest rate of monetary policy on economic output. When People's Bank of China lowers the MLF rate, the LPR goes down and affects the total output of the economy through the response of banks, businesses, and residents and then gradually returns to the vicinity of steady state. Taking into account the uncertainty of COVID-19 in 2020–2021, this is a comparative analysis of the 2019 national economic statistics before the epidemic: in the year LPR (1 year) dropped 3 times, with a year-end value of 4.15%, which was down by 26BP compared with the beginning of the year. The LPR (5-year period) was set up in August 2019, 1 time lower, and the value at the end of the year was 4.80%, which decreased by 5BP compared with the initial value. The annual price index is 102.90%, with an increase of 0.2 percentage point from the previous year. The proportion of RMB loans balance as a percentage of GDP is 154.52% and the proportion of RMB deposit balance as a percentage of GDP is 97.10%. The new RMB loans account for 66.0% of the new social financing scale in the same year. The year-on-year growth rate of the money supply M2 balance is 8.70%, with an increase of 0.1 percentage point from the previous year. The NIM of commercial banks is 2.19%, with an increase of 0.14 percentage point compared with the end of the previous year (Table 3).

This paper analyzes the matching degree between benchmark model simulation results and China's macro-economic data in 2019 in terms of interest rate level, bank scale, and the real economy.

TABLE 2: Calibration values of main parameters.

Interest rate level	Value	Parameter	Value	Parameter	Value
$\beta^t$	0.988	$T_p$	0.412	$\varphi_\pi$	0.023
$\sigma_c$	0.331	$\varepsilon$	0.201	$\varphi_y$	0.921
$\sigma_n$	0.355	$\beta_i^p$	0.122	$\varphi_m$	0.634
$\sigma_m$	0.423	$r_i$	0.311	$\varphi_p$	0.123
$\alpha_1$	0.562	$\theta_R$	0.483	$m_p$	0.913

In terms of policy rate level, the 1-year MLF interest rate (3.35%) and LPR interest rate (4.90%) simulated by the model are both higher than the corresponding actual interest rate levels. The simulated loan interest rate (5.25%), deposit interest rate (2.80%), and government bond interest rate (2.70%) are all lower than the corresponding actual interest rate levels. This indicates that the monetary policy transmission mechanism in China still has a serious obstacle in practice. The prudent risk behavior of the banking system in the credit process makes its loans with a relatively long period and large loan limit and concentration; i.e., the actual loan interest rate corresponding to the policy rate is higher than the equilibrium interest rate level simulated by the model. The actual deposit interest rate corresponding to the policy rate is also higher than the equilibrium interest rate simulated by the model due to the competition for residents' deposits and the substitution of fixed time and current deposits by certificate deposits and finance products.

In terms of bank scale, the proportions of deposit and loan balances as a percentage of GDP simulated by the model are highly consistent with the actual data. However, the proportion of new loans as a percentage of new social financing is lower than the actual value, and the proportion of new government bonds as a percentage of new social financing is higher than the actual value, which indicates that China's financial market structure is still unsatisfactory, with a high proportion of credit financing and a relatively low proportion of capital market financing. The net interest margin enjoyed by Chinese commercial banks is still higher than the equilibrium value due to the blockage of policy rate transmission, indicating that the policy dividends still exist.

TABLE 3: Comparison of model simulation results with actual data of China's economy (2019).

Interest rate level	$R_t$	LPR <sub>t</sub>	$R_t^P$	$r_t$	$R_t^G$
Model	3.35	4.90	5.25	2.80	2.70
Data	3.26	4.80	5.50	3.80	2.85
Bank scale	Deposit/GDP	Loans/GDP	New loans/new social financing	New government bonds/new social financing	Net interest margin
Model	96.7%	155.1%	60%	20%	2.05%
Data	97.1%	154.5%	66%	18.5%	2.19%
Real economy	GDP growth rate	Inflation rate	Unemployment rate	M2 growth rate	Enterprise asset-liability ratio
Model	6.3%	2.7%	4.3%	8.6%	62%
Data	6.1%	2.9%	4.5%	8.7%	64%

In terms of the development of the real economy, the GDP growth rate, inflation rate, unemployment rate, M2 growth rate, and enterprise asset-liability ratio simulated by the model are all superior to the actual data, which indicates that the blockage of policy rate transmission has aggravated the business difficulties of the real economy and reduced the operating efficiency of China's economy.

## 5. Results and Discussion

This paper used the Dynare tool to estimate the dynamic response of each variable in the economy to the temporary impact of the given policy rate with the advancement of the LPR interest rate reform. Specifically, it can be divided into the following scenarios: (i) LPR pricing mechanism loans account for 50% of the existing loans; (ii) LPR pricing mechanism loans account for 80% of the existing loans; (iii) LPR pricing mechanism loans account for 100% of the existing loans. This paper takes the weighted average interest rate of one-year loan, the interest rate of one-year time deposit, and the yield of one-year treasury bond as observation variables.

From August 2019 to the end of 2021, in order to hedge the impact of COVID-19, China's monetary policy was in an easing cycle. The 1-year LPR interest rate has been lowered by 6 times and the cumulative 35BP has been lowered.

*5.1. LPR Pricing Mechanism Loans Account for 50% of the Existing Loans.* In this scenario, we assume that there are no other changes in policy constraints. The model results showed that when the policy rate decreased by 35 BP, the loan interest rate, deposit interest rate, and government bond interest rate declined by 21 BP, 15 BP, and 26 BP. This indicates that when LPR pricing mechanism loans account for 50% of the existing loans, the transmission efficiencies of the policy rate to the loan interest rate, deposit interest rate, and government bond interest rate are 60%, 43%, and 74%, respectively (Figure 2).

Generally speaking, when the policy rate falls, the price of borrowing funds from the central bank by commercial banks decreases, and commercial banks will reduce their demand for deposits, thus resulting in a decline in deposit interest rate, which in turn leads to a diversion of residents' deposits to Internet finance as well

as the stock and bond markets. In practice, however, the falling range of deposit interest rate will be smaller because the deposit interest rate is relatively stable, while commercial banks, due to the price reduction of funding sources, will cut enterprise loan interest rate accordingly in order to gain more loan market shares, thus bringing down the enterprise financing cost. Accordingly, the price of bond financing in the financial market will also decline. Therefore, from the model, when the policy rate decreases, the loan scale of commercial banks will increase, while the deposit scale will decline.

*5.2. LPR Pricing Mechanism Loans Account for 80% of the Existing Loans.* Similarly, in this scenario, we assume that there are no other changes in policy constraints. The model results showed that when the policy rate decreased by 35 BP, the loan interest rate, deposit interest rate, and government bond interest rate declined by 24 BP, 18 BP, and 29 BP, respectively. Thus it is indicated that when LPR pricing mechanism loans account for 80% of the existing loans, the transmission efficiencies of the policy rate to the loan interest rate, deposit interest rate, and policy bond interest rate are 69%, 51%, and 83%, respectively.

*5.3. When LPR Pricing Mechanism Loans Account for 100% of the Existing Loans, The Benchmark Interest Rates of Deposits and Loans Are Removed.* Similarly, in this scenario, we assume that there are no other changes in policy constraints. The model results showed that when the policy rate decreased by 35 BP, the loan interest rate, deposit interest rate, and government bond interest rate declined by 28 BP, 25 BP, and 32 BP, respectively. Thus it is indicated that when LPR pricing mechanism loans account for 100% of the existing loans, the transmission efficiencies of the policy rate to the loan interest rate, deposit interest rate, and policy bond interest rate are 80%, 71%, and 91%, respectively.

*5.4. The Impact Effect of a Unit Interest Rate Change on Economic Output under Three Different Scenarios.* Under the above three different marketization degrees, the adjustment of LPR interest rate can have an impact on the actual output. Further observing the impulse intensity under different

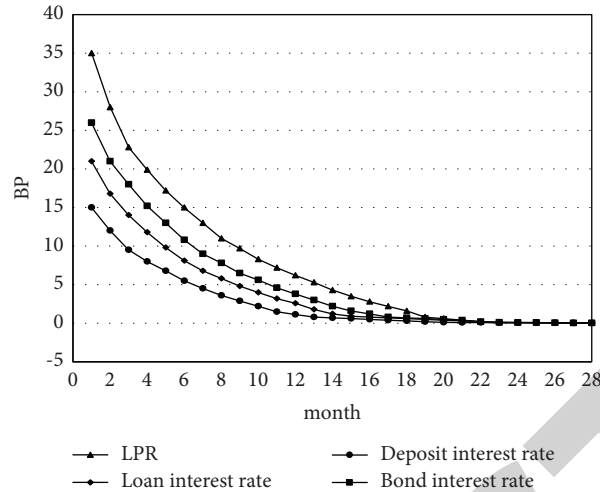


FIGURE 2: Effect of LPR interest rate on deposit, loan, and treasury bond interest rate.

marketization degrees, we can find that, with the continuous relaxation of interest rate control, output is becoming more sensitive to interest rate adjustments. From the perspective of transmission mechanism, the reduction of LPR interest rate means the decline of policy interest rate, which is a loose monetary policy and can effectively promote economic growth. This paper uses the year-on-year growth of real GDP as the measurement index of economic output and uses Dynare tool to estimate the impact of economic output on the temporary impact of reducing one unit of LPR interest rate. The following results can be obtained: during the period when the LPR pricing mechanism loans account for 50% of the loan stock, reducing the LPR interest rate by 1 BP will increase the year-on-year growth rate of real GDP by 0.005 percentage points; in the period when the LPR pricing mechanism loan accounts for 80% of the loan stock, reducing the LPR interest rate by 1 BP will increase the year-on-year growth rate of real GDP by 0.009 percentage points; in the period when the LPR pricing mechanism loan accounts for 100% of the loan stock, reducing the LPR interest rate by 1 BP will increase the year-on-year growth rate of real GDP by 0.012 percentage points, as shown in Figure 3.

The results show that the LPR pricing system reform implemented by People's Bank of China makes the interest rate marketization enter a new stage, can effectively improve the interest rate transmission efficiency of monetary policy, and plays a positive role in promoting economic growth. First, it really decouples the loan pricing of commercial banks from the policy benchmark interest rate and promotes the LPR interest rate to 100% existing loan pricing, indicating the substantive progress in the market-oriented reform of interest rate; second, it adopts the gradual principle to promote the application of LPR interest rate and takes MLF interest rate as the reference standard to prevent the excessive fluctuation of market interest rate to a certain extent; third, LPR interest rate has a certain impact on real GDP growth, and the impact effects are quite different under different marketization degrees, which means that financial market friction has a certain obstacle to the transmission

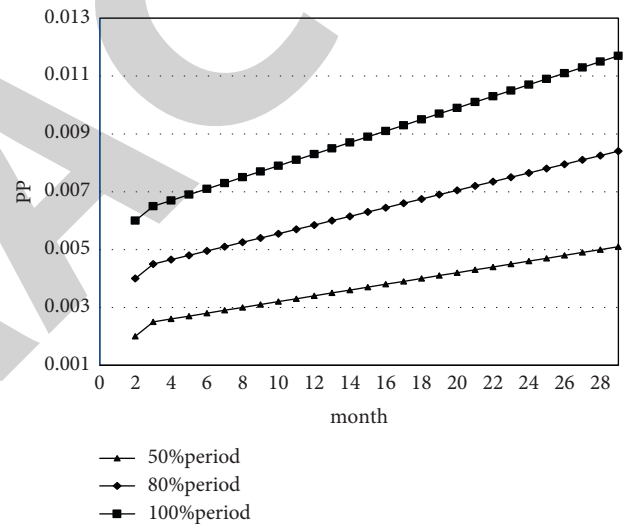


FIGURE 3: Impulse effect of LPR interest rate on economic growth under different degrees of marketization.

efficiency of monetary policy. Since LPR not only inherits the advantages of gradual reform, but also integrates market-oriented elements, the interest rate transmission mechanism after its emergence needs to be researched. In addition, it is worth noting that the monetary transmission efficiency in the DSGE theoretical model is generally better than the actual data. This means that the efficiency loss faced by monetary policy in the actual transmission process during the regulation period is much greater than the theoretical expectation. At the same time, it also shows that LPR, an important reform that takes into account both robustness and market attributes, will become a useful attempt in the process of interest rate marketization.

## 6. Research Conclusions

This paper constructs a DSGE model to compare and analyze the impact of LPR reform on monetary policy transmission mechanism and economic output under different

marketization degrees and mainly draws the following conclusions. First, interest rate regulation has a real loss on the transmission efficiency of monetary policy, which is far greater than the theoretical expectation; second, with the advance of LPR reform, the response of real GDP year-on-year growth to LPR interest rate has become more and more sensitive, indicating that LPR interest rate reform can effectively affect economic growth; third, the higher the degree of marketization of LPR interest rate is, the more significant the transmission effect on deposit interest rate, loan interest rate, and treasury bond interest rate will be, and the efficiency of interest rate transmission mechanism will be improved. Consider that LPR pricing mechanism is still in the process of improvement, and in the meantime, there is a negative impact of COVID-19 on China's economy. The empirical results of this paper need to be further improved and verified.

At the same time, we still need to deeply realize that there are still some differences between the theoretical expectation and practical effect of the reform. Especially in the stage of complete marketization, the actual efficiency loss of progressive reform is much higher than the theoretical expectation. For example, at the beginning of implementing the interest rate corridor model by the European Central Bank, the width of the interest rate corridor reached 250 BP. After 2003, in the face of the chain impact of the subprime mortgage crisis of the US, the European Central Bank continuously narrowed the interest rate corridor, of which the interest rate corridor was narrowed to 100 BP in December 2008 and 50 BP in June 2014 and is now maintained at 65 BP. This has led to a sharp narrowing of the interest spreads of deposit and loan of commercial banks, and some banks have thus got into trouble. This means that there are many hidden costs in the regulated interest rate transmission mechanism. At the same time, it also shows that only by fundamentally promoting the complete market-oriented reform will the interest rate transmission efficiency be fundamentally changed.

It is suggested that People's Bank of China continues to strengthen its research on improving the marketization of interest rates: First, a more complete interest rate corridor mechanism should be created, and the interest rate corridor range should be gradually narrowed, so as to ensure to give full play the role of the interest rate corridor mechanism. Second, the monetary policy transmission mechanism should be improved, and the function of the MLF as a medium-term policy rate should be exerted while strengthening the guidance of short-term policy rate. Third, the benchmark interest rates of loans and deposits determined by the central bank should be removed appropriately. Fourth, the direct financing channels should be vigorously developed and the monetary policy transmission mechanism should be improved. In this process, as the LPR reform is a gradual implementation process, the depth and breadth of its application have an important impact on the transmission of policy rate to market rate.

It is suggested that the commercial banks actively adapt to the process of interest rate marketization reform. First, it is necessary to establish a more scientific and reasonable

pricing mechanism and improve the depth of application of LPR. Second, it is necessary to enhance the foresight of interest rate risk management, establish a mechanism for hedging interest rate exposures, and reasonably use interest rate swaps and other derivative instruments to carry out hedging operations. Thirdly, the innovation of interest rate derivative products can ensure the effect of interest rate market reform and further unblock the efficiency of monetary policy transmission.

## Data Availability

The data were taken from the WIND database and relevant information from People's Bank of China. Due to confidentiality requirements, they cannot be provided to research institutions outside of mainland China.

## Conflicts of Interest

The author declares that there are no conflicts of interest.

## Acknowledgments





This work was supported by Beijing Municipal Social Science Foundation (Grant no. 20JCB070).

## References

- [1] J. Tobin, "A general equilibrium approach to monetary theory," *Journal of Money, Credit, and Banking*, vol. 1, no. 1, pp. 15–29, 1969.
- [2] F. Modigliani, "Monetary policy and consumption: linkages via interest rate and wealth effects in the FMP model," in *Consumer Spending and Monetary Policy: The Linkages, Conference Series*, pp. 5–15, no. 4, Federal Reserve Bank of Boston, Boston, MA, USA, 1971.
- [3] B. S. Bernanke and A. S. Blinder, "The federal funds rate and the channels of monetary transmission," *American Economic Review*, vol. 82, pp. 901–921, 1992.
- [4] B. S. Bernanke and M. Gertler, "Inside the black box: the credit channel of monetary policy transmission," *NBER Working Paper*, vol. 9, no. 2, pp. 25–30, 1996.
- [5] F. S. Mishkin, "The channels of monetary transmission: lessons for monetary policy," *NBER Working Paper*, pp. 2–27, 1996.
- [6] C. L. Evans and D. A. Marshall, "Monetary policy and the term structure of nominal interest rates: e," in *Carnegie-Rochester Conference Series On Public Policy*, vol. 49, pp. 53–111, North-Holland, 1998.
- [7] G. J. De Bundt, "Interest rate pass-through: empirical results for the euro area," *German Economic Review*, vol. 6, no. 1, pp. 37–38, 2005.
- [8] P. Bolton and X. Freixas, "Corporate finance and the monetary transmission mechanism," *Review of Financial Studies*, vol. 19, no. 3, pp. 829–870, 2006.
- [9] F. Milani and J. Treadwell, "The effects of monetary policy "news" and "surprises"," *Journal of Money, Credit, and Banking*, vol. 44, no. 8, pp. 1667–1692, 2012.
- [10] S. Gomes, N. Iskrev, and C. Mendicino, "Monetary policy shocks: we got news!" *Journal of Economic Dynamics and Control*, vol. 74, no. 1, pp. 108–128, 2017.

## Research Article

# Edge Caching in Fog-Based Sensor Networks through Deep Learning-Associated Quantum Computing Framework

Tayyabah Hasan,<sup>1</sup> Fahad Ahmad ,<sup>2</sup> Muhammad Rizwan ,<sup>1</sup> Nasser Alshammari ,<sup>3</sup> Saad Awadh Alanazi ,<sup>3</sup> Iftikhar Hussain,<sup>4</sup> and Shahid Naseem<sup>5</sup>

<sup>1</sup>Department of Computer Sciences, Kinnaird College for Women, Lahore 54700, Punjab, Pakistan

<sup>2</sup>Department of Basic Sciences, Deanship of Common First Year, Jouf University, Sakaka 72341, Aljouf, Saudi Arabia

<sup>3</sup>Department of Computer Science, College of Computer and Information Sciences, Jouf University, Sakaka 72341, Aljouf, Saudi Arabia

<sup>4</sup>Division of Engineering Management and Decision Sciences, College of Science and Engineering, Hamad Bin Khalifa University, Doha 34110, Qatar

<sup>5</sup>Department of Information Sciences, Division of Sciences and Technology, University of Education, Lahore 54770, Pakistan

Correspondence should be addressed to Fahad Ahmad; drfahadahmadmian@gmail.com

Received 1 October 2021; Revised 1 November 2021; Accepted 3 November 2021; Published 7 January 2022

Academic Editor: Daqing Gong

Copyright © 2022 Tayyabah Hasan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Fog computing (FC) based sensor networks have emerged as a propitious archetype for next-generation wireless communication technology with caching, communication, and storage capacity services in the edge. Mobile edge computing (MEC) is a new era of digital communication and has a rising demand for intelligent devices and applications. It faces performance deterioration and quality of service (QoS) degradation problems, especially in the Internet of Things (IoT) based scenarios. Therefore, existing caching strategies need to be enhanced to augment the cache hit ratio and manage the limited storage to accelerate content deliveries. Alternatively, quantum computing (QC) appears to be a prospect of more or less every typical computing problem. The framework is basically a merger of a deep learning (DL) agent deployed at the network edge with a quantum memory module (QMM). Firstly, the DL agent prioritizes caching contents via self organizing maps (SOMs) algorithm, and secondly, the prioritized contents are stored in QMM using a Two-Level Spin Quantum Phenomenon (TLSQP). After selecting the most appropriate lattice map ( $32 \times 32$ ) in 750,000 iterations using SOMs, the data points below the dark blue region are mapped onto the data frame to get the videos. These videos are considered a high priority for trending according to the input parameters provided in the dataset. Similarly, the light-blue color region is also mapped to get medium-prioritized content. After the SOMs algorithm's training, the topographic error (TE) value together with quantization error (QE) value (i.e., 0.0000235) plotted the most appropriate map after 750,000 iterations. In addition, the power of QC is due to the inherent quantum parallelism (QP) associated with the superposition and entanglement principles. A quantum computer taking " $n$ " qubits that can be stored and execute  $2^n$  presumable combinations of qubits simultaneously reduces the utilization of resources compared to conventional computing. It can be analyzed that the cache hit ratio will be improved by ranking the content, removing redundant and least important content, storing the content having high and medium prioritization using QP efficiently, and delivering precise results. The experiments for content prioritization are conducted using Google Colab, and IBM's Quantum Experience is considered to simulate the quantum phenomena.

## 1. Introduction

Fog computing (FC), at the edge of a sensor network, as an extension to cloud computing, offers storage, processing, and communication control services [1, 2]. In the period of

next-generation telecommunication and through the massive development of the Internet of Things (IoT) based smart devices, applications required ultralow latency because IoT networks induce strain not only on the backhaul but the fronthaul causing adverse situations for interruption



sensitive applications [3, 4]. These problems can be resolved through FC, which provides distributed computing and communication facilities from centralized servers in the edge direction. A central base band unit (BBU) pool is not robust for every control, communication, or any other processing function; therefore, FC-based radio access networks (F-RANs) were introduced. In F-RANs, the local BBUs or even remote radio heads (RRHs) are dedicated to such tasks through edge caching (EC) [5–9]. Due to an intermediate fog layer between end-users and the cloud, mobile edge computing (MEC) is introduced.

Although the idea of F-RANs seems to be propitious to provide all the tasks confronted by the cloud radio access networks (CRANs) or heterogeneous cloud radio access networks (HCRANs). However, some setbacks might cause performance deterioration or the quality of service (QoS) degradation, bringing about fronthaul congestion. The main issue that requires to be solved is EC as well as restricted storing capability in RRHs [10]. Limited vital interests associated with the EC trending in F-RANs are reducing fronthaul burden, backhaul, or even backbone, optimizing endwise latency issues, and dynamic applications of content responsive approaches performance improvements. Fog access points (FAPs) usually have a relatively minimal caching capacity mostly because of the limited memory linked to the caching processes executed in centralized CRANs or HCRANs. Nevertheless, an increase in cache size in the base stations (BS) has a balance in the middle of improved throughput and network spectral efficiency [11]. Consequently, caching techniques in FAPs, together with caching strategies and allocation of cache resources policies, need to be managed logically and dynamically for augmented F-RAN performance.

The main contribution of this research is to predict the high priority content through the deep learning (DL) technique. It is the leading task that must be carried out when the contents are requested repeatedly and placed in the caches. The rest of the content should be discarded. When the high priority content is predicted through the DL agent, efficient content management and placement are achieved through the proposed framework and the quantum memory modules (QMM) to store the content. This paper describes an EC-based deep learning-associated quantum computing (DLAQC) framework. The framework is based on two parts: one for caching content prioritization and the other one for caching content stored within the edge. The DL-based quantum computing (QC) approach associated with quantum information processing is deployed to enhance the performance of F-RANs. The framework is basically a merger of a DL agent deployed at the network edge and a QMM. Firstly, the DL agent prioritizes caching contents via Self-Organizing Maps (SOMs) algorithm, and secondly, the prioritized contents are stored in QMM using a Two-Level Spin Quantum Phenomenon (TLSQP). SOMs algorithm is staunchly suitable to pick up contents in colored cluster form without reducing the dimensionality of the feature space.

The paper's organization is as follows. Section 2 describes the related work for edge caching, SOMs applications, and Stern–Gerlach experiment (SGE). In Section 3, the

framework and algorithm are described, followed by an overview of the model. The DL agent in edge and TLSQP is also discussed. The experiments and results are analyzed and discussed in Section 4. Finally, conclusion is presented in Section 5.

## 2. Literature Review

In this section, a literature study is carried out to throw light on attempts of different researchers to enhance EC to improve efficiency and quality of service (QoS) in F-RANs. Several pieces of research related to the DL-based algorithm and SGE highlight their applications in various fields, which has proved to be a great source of guidance for the proposed idea.

The authors in [12] described the key features of MEC, especially in the context of next-generation and IoT-based applications. The role of MEC and its challenges in the context of edge intelligence is also described. By keeping in view, the latency, context awareness, and energy-saving criteria, it is compared with the conventional MCC by considering the following key enabling features: virtual reality/augmented reality, software defined network, network function virtualization, smart devices, information-centric networking, network slicing, and computation off-loading. Additionally, a use case is also provided to help understand the edge intelligence in the IoT-based scenarios.

The critical challenges in the F-RANs are (i) the content placement in caches and (ii) the joint user associations. These challenges are tinted due to the complexity of different approaches used to find optimal solutions. In [13], authors considered optimization problems as mixed-integer non-linear programming (MINL). A hierarchical game theory approach is applied, and a series of deep reinforcement learning (DRL) based algorithms are designed for user association, content popularity prediction, and content placement to enhance the FAPs. In [14], a cooperative EC scheme using the DRL approach to place and deliver contents in vehicular edge computing networks is presented. The deep deterministic policy gradient algorithm provides a sensible solution for long-term MINL problems. A scheme for vehicle scheduling and bandwidth allocation is designed to make it less complex to manage adaptive resources and make decisions.

A user preference-based learning EC policy is described in [15] to predict the online content popularity and an offline learning algorithm. A sigmoid function is exploited to construct a logistic regression model to estimate user preferences regarding online content popularity prediction. It is considered complicated to make a preference model for each user due to the high-dimensional feature space. Therefore, a follow-the-regularized-leader proximal inspired algorithm is also proposed for offline user preference learning.

The federated learning (FL) framework and deep reinforcement learning (DRL) techniques were integrated into MEC in [16] to optimize EC, computing, and communication. The key challenge is primarily faced by authors to place the DRL agents due to (i) the massive and redundant

data transmissions in the cloud and (ii) the privacy risks as well as the lesser computing capabilities in UEs. The proposed technique has outperformed the conventional caching policies for EC like Least Recently Used, Least Frequently Used, and First In First Out. However, for computation offloading, the DRL technique by FL offered close results to the centralized DRL technique; on the other hand, again, it outperformed the following baseline policies: mobile execution, edge node execution, and greedy execution.

The authors in [17] have presented an on-demand and collaborative deep neural networks (DNN) coinference framework. The presented framework worked in two ways. Firstly, the DNN computation is partitioned between devices and the edge to make use of hybrid computation resources, and it can exit early in the DNN right-sizing at some suitable intermediate DNN layer. Therefore, it can avoid further computation latency and have implemented their prototype on Raspberry Pi. Secondly, the visualization of higher-dimensional data is taken into account to effectively analyze and conclude the data and results. The following two strategies have been used to visualize the multidimensional and minimal data by a scatter plot established on dimensionality reduction: (i) the direct visualization and (ii) the projection methods.

Failure modes and effects analysis (FMEA) is a methodology for risk analysis and problem prevention by identifying and defining failures of the system, process, or service. FMEA has some shortcomings related to the worksheet and usage complexity, which have been dealt with by the SOMs algorithm in [18]. SOMs algorithm is exploited to achieve perceptibility for corrective actions. A risk priority interval is used to evaluate these corrective actions in groups to make it easier for the users.

SOMs for multiple travelling salesman problem (MTSP) with minmax objective is exploited for the robotic multigoal path planning problem [19]. The main issue in deploying this framework was to detect the collision-free paths to evaluate the distances in the winner selection phase. The collision-free path was needed to adapt the neurons to the presented input signals. To address this issue, simple approximations of the shortest path are considered and verified through cooperative inspection. The presented SOMs approach is used to solve this inspection task by MTSP-minmax and compared with the MTSP-GENIUS algorithm.

In [20], the SOMs have been used to classify astronomical objects like stars' stellar spectra. The algorithm is used to make different spectral classes of the Jacoby, Hunter, and Christian library. The 158 spectra were chosen to classify by 2799 data points each. 7 clusters were formed from O to M, and 12 out of 158 spectra were misclassified, giving a 92.4% success rate. Otto Stern and Walther Gerlach, in 1922, performed an experiment that separated an electron beam while passing through a nonuniform magnetic field. When a beam passed through a magnetic field, two distinguished beams were observed on the screen. The experiments conducted by Otto Stern and Walther Gerlach gained popularity and were used in multidisciplinary studies by researchers.

In [21], the SGE is exploited in physical chemistry to investigate the magnetic response of the Fe@Sn12 cluster. A comparison is carried out between Mn@Sn12 and Fe@Sn12 clusters by passing their beams through magnetic fields separately. The molecular beam of Fe@Sn12 cluster exclusively deviates more towards increasing the magnetic field. The beam deviates even at the shallow temperature due to the distortions of tin-cage induced by Jahn-Teller. Hence, in the magnetic dipole moment, the role of electronic orbital angular momentum is significant. The magnitude of the magnetic dipole moment is calculated from the transfer of the beam.

In [22], the SGE is oppressed to explore the spin  $\frac{1}{2}$  neutral particles' motion and how their motion is dependent on the initial phase difference between two internal spin states. If the particles are moving horizontally, the initial phase difference between spin states results in particle splitting in the longitudinal direction and in the lateral direction due to the quantum interference. This interference provides an alternate way of measuring the initial phase difference between spin states and helps determine the amplitude and phase of atoms in the same SGE. To study this phenomenon, the ultracold temperature is maintained to make the ideal condition for the atom to behave like a quantum wave packet instead of a particle. In general, an atom is not in a pure state, rather a mixed state and cannot be characterized as a single wave function.

The content priority is deduced by the adaptive neurofuzzy inference system (ANFIS) in [23] in which the following five input variables were carefully selected: video\_elapsed\_time, video\_size, views, likes, and downloads. Each input variable has three membership functions having priorities high, medium, and low, and fifteen similar functions are made in the Sugeno inferencing mode. A rule base function was also created. After content prioritization through ANFIS, a theoretical explanation of the SGE is specified as a TLSQP for storing the prioritized content in quantum repositories.

Considering media requirements explicitly, EC seems to alleviate certain challenges. Occasionally, multimedia contents get heavier than even the encyclopedias, resulting in higher hardware and network capacities. EC can support such kinds of throughput requirements proportionally. Moreover, the scalability of streaming servers, which require special provisioning of these servers, can also be handled during live events. However, reducing the distance between end devices and BS will not be sufficient for increased network throughput; high-speed backhaul connectivity is also required between all the BS and the BBU Pool where centralized servers reside. The network traffic load can be reduced by minimizing redundant traffic. The traffic load mainly comprises content deliveries for the most requested/popular content at different times. If at all this redundant traffic is managed so that popular content is predicted and placed within the edge, the idea's effectiveness to increase network efficiency can be justified. Researchers and network specialists also have incorporated different AI techniques, including machine learning, to minimize the redundancy of network traffic and optimize the overall network efficiency,

from predicting popular content to optimizing specific parameters and much more [24–27]. In this digital era, IoT-based devices have generated an enormous amount of data daily, which is one reason for the possible growth of DL algorithms [28]. The DL algorithms require a massive amount of data to learn from.

### 3. Deep Learning-Associated Quantum Computing Framework

Edge computing acts as an intermediary between cloud and user equipment through the network edge. Researchers and engineers are continuously trying to accelerate content deliveries further and make mobile services better. The intelligence of edge systems is enhanced by introducing a DL agent in network edge. The DL agent is used to prioritize (i) the caching content according to its popularity determined by the considered parameters and (ii) the content to be managed logically within the planned framework. Prioritizing the content intelligently for caching only is not adequate to optimize the overall performance of the system. The prioritized contents need to be stored efficiently in caches and accessed multiple times with instantaneous delivery response. Within edge caches, the TLSQP will take over to avoid limited storage issues.

*3.1. Overview.* To understand the workflow, a model is presented wherein the fog environment is described together with the proposed DLAQC framework and shown in Figure 1.

A particular region is considered from where user requests are generated. The cloud servers initialize the fog environment’s monitoring cycle from the BBU pool. In the fog layer, at every moment, numerous user requests are generated and served through the F-APs. In order to accelerate content deliveries or response time, a DLAQC framework is presented, and a brief overview is as follows:

- (1) Synchronizer: cache synchronization and inter-F-APs information sharing is carried out every moment to update the regional user set for particular F-APs in the region
- (2) Regional user set: it refers to the group of users from a particular region allotted to a locally installed F-AP for a particular period.
- (3) Local and neighboring F-APs: the F-APs are capable of caching and computation and serve user requests by searching through the caches within the edge. A particular user request is immediately served if the content is available locally. The content must be looked up from the neighboring F-APs caches when a local cache is missed. In case if the contents are not available in the neighboring F-APs, the offered framework will update the respective contents in caches’ QMM.
- (4) DL agent for content prioritization: in case of a cache miss, the content is intelligently updated through the DL agent deployed at the edge. It comprises SOMs and helps to

predict the contents’ priorities. The contents having maximum demand are considered highly prioritized. However, this module is used to logically manage the contents (specifically that need to be updated).

- (5) Quantum memory module: it is one of the most critical modules in the given setup. Once the contents are prioritized intelligently through the DL agent, the contents need to be stored optimally in caches to enable caching content management. The synchronizer module is used to do so. The requested content may also be served through (i) F-APs located in the same region or (ii) user dynamics or load balancing. The QMM is incorporated especially to place contents physically in caches as quantum particles when it is prioritized. As a case, in this model, Repository 1 is assigned for storing highly prioritized contents, and Repository 2 is assigned for the medium prioritized contents, respectively. QMM is used to store and serve (the requested) contents separately. Due to the lower demand, every time, the low-priority contents are discarded from the caches.

The proposed framework comprises two modules: DL agent and QMM. The DL agent is used to prioritize and logically manage the caching contents by making use of SOMs. The QMM is used to store the contents in a quantum regime by exploiting a TLSQP. The proposed framework’s problem (working) and solution (implementation) domains are described as follows.

#### 3.2. Framework (Problem)

*3.2.1. Deep Learning Agent in Edge.* The DL agent is deployed at the network edge and prioritizes the caching contents through SOMs [29]. It gives results similar to the clustering approaches, and the prioritized contents can be visualized through light and dark color concentrations. The graphical output given by SOMs is a kind of feature map for input space. It makes SOM suitable for prioritizing the content using specific parameters. In this study, the media contents are explicitly targeted. Dataset for Trending YouTube Videos Statistics has been downloaded from Kaggle. The dataset includes the statistics for trending videos in the region of the United States. To achieve the requirements, four input variables for each video are carefully selected from the dataset and which are as follows: views, likes, dislikes, and comment\_count. The identified relationship between input parameters is helpful for visualizing the trending contents’ priorities using SOMs. The structure and function of SOMs are explained by the mathematical model as follows. SOMs work by fitting the map (grid of nodes) up to the given number of iterations of the simulated dataset. During diverse iterations, the adjustments are required while nodes’ weights are adjusted to bring the map nodes closer to the data points. It is called the convergence of SOMs, and the structure for SOMs is given in Figure 2.

The main package is included to construct, evaluate, and visualize the map is Minisom. An input layer (4-dimensional) and feature space  $\mathbf{M}$  of the map are defined by the

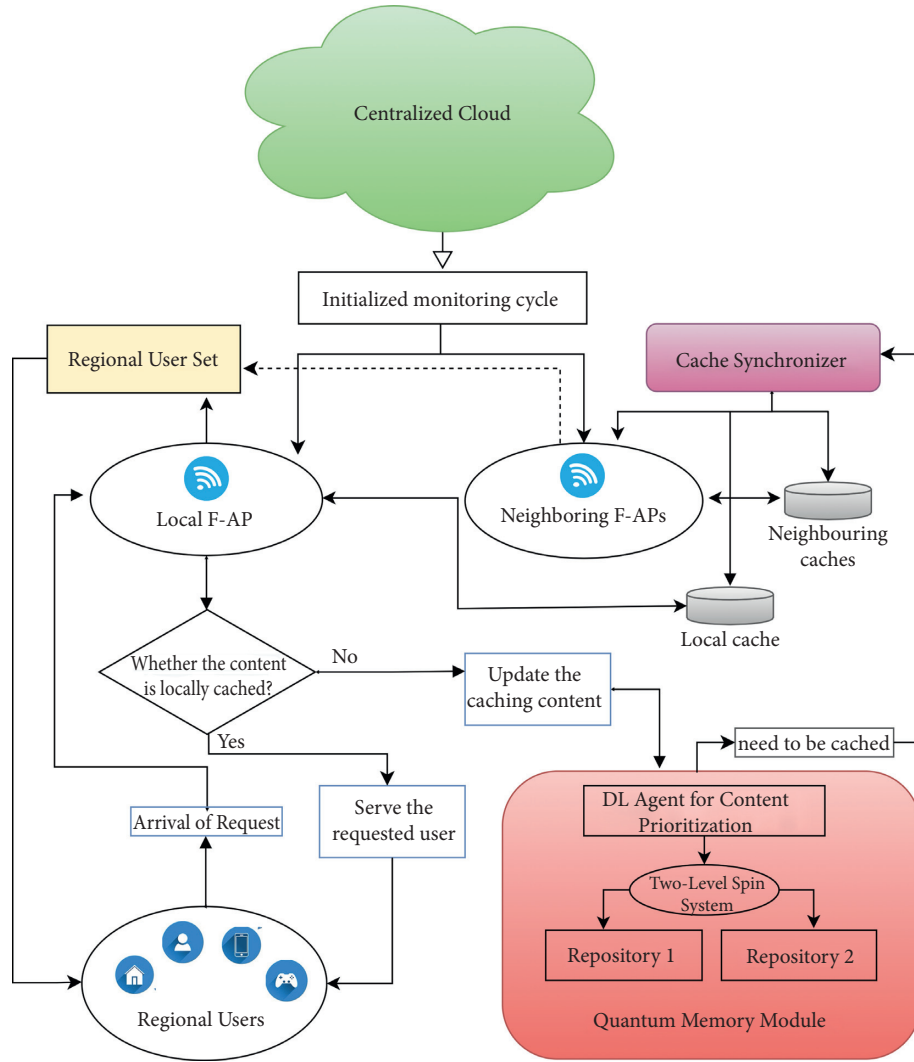


FIGURE 1: The proposed deep learning-associated quantum computing (DLAQC) framework.

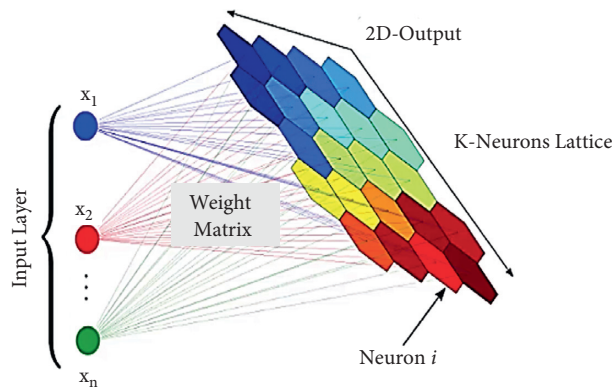


FIGURE 2: Structure of Self-Organizing Maps.

rule of thumb. The rule of thumb states that there should be  $5 \cdot \sqrt{N}$  neurons in the lattice to get desired results, where  $N$  is the total number of samples in the dataset (the training dataset). The training dataset has 40960 samples; thus, the lattice should contain  $5(\sqrt{40949}) = 1011.8$  neurons.

Therefore, the dimensions of the lattice are selected as  $32 \times 32$ . Each node in the lattice has a weight vector  $W_{ij}$  and has the same dimensions as input vectors  $V$ . The preliminary step of training is to set weights of every node and is initialized as  $W_{ia} : W_{ib} : W_{ic} : W_{id}$  where  $i$  represents node



number and  $\mathbf{a}$ ,  $\mathbf{b}$ ,  $\mathbf{c}$ , and  $\mathbf{d}$  represent input vectors. When the weights are initialized, the best matching unit BMU is calculated by iterating through every node and by calculating the Euclidean distance between input vector  $\mathbf{V}$  and each node's weight  $\mathbf{W}$ . Finally, the smallest  $\mathbf{W}$  is selected. The process is given as follows:

$$\mathbf{D} = \sqrt{\sum_{i=0}^n (\mathbf{V}_i - \mathbf{W}_i)^2}. \quad (1)$$

When the BMU is finalized, the neighborhood nodes whose weights need to be updated are determined. To achieve this, the Gaussian neighborhood function is used. In this function, the "bell-shaped curve" like weighting is considered to update the nodes depending upon their relative distances from the BMU. Initially, the sigma  $\sigma_0$  is used to denote the spread of the neighborhood function, and all nodes (come in this spread) are updated. The respective spread shrinks iteratively by using the function (decay function) described as follows.

$$\sigma_t = \frac{\sigma_0}{(1 + t)/(T/2)}, \quad (2)$$

where  $\mathbf{T}$  is the iterations set having  $\{t_0, t_1, t_2, t_3, \dots, t_n\}$  and  $\sigma_t$  is the spread size at iteration  $t$ . Every node in the neighborhood of BMU is updated by

$$\mathbf{W}_{(t+1)} = \mathbf{W}_t + \Theta_t [\mathbf{L}_t (\mathbf{V}_t - \mathbf{W}_t)]. \quad (3)$$

A decaying function of learning is given as follows, where  $\mathbf{L}_t$  is the learning rate.

$$\mathbf{L}_t = \frac{\mathbf{L}_0}{(1 + t)/(T/2)}. \quad (4)$$

Moreover,  $\Theta_t$  is the distance effect from the BMU on the specific node and is given as follows:

$$\Theta_t = e^{-[D^2/2\sigma^2]}. \quad (5)$$

Hence, blocks with similar color zones are visualized. Any new input will stimulate nodes in the zone with similar weight vectors. The process described above results in projection of all the data points onto the map that allows topology of high-dimensional input data to be preserved into two-dimensional output space. However, the visual inspection is not enough to determine (i) how well the map converges to the given data points or (ii) how well the map represents the underlying data. Some quality measures are developed to oblige the purpose of evaluating when the map is trained. Therefore, the Quantized Error (QE) is used for vector quantization to evaluate the quality of the map. It is achieved by summing up the distances between the nodes and the data points as per the average distance given as follows.

$$\text{QE}(\mathbf{M}) = \frac{1}{\mathbf{N}} \sum_{i=n}^n \|\varphi(\mathbf{x}_i) - \mathbf{x}_i\|, \quad (6)$$

where the feature space of the map is denoted as  $\mathbf{M}$ .  $\mathbf{N}$  is used to represent the total number of data points and  $\varphi(\mathbf{x}_i)$  is used for mapping of data point  $\mathbf{x}_i$  from input space to the map.

Hence, it is considered as; the smaller the value of QE, the better it fits the data points. However, this quality measure can be used to compare maps by considering the same dataset and choosing the best one, not as the only quality assessment.

One of the primary aims of SOMs to determine quality is the topological preservation of high-dimensional input space in the two-dimensional output space. The topographic error (TE) is used to evaluate how well the individual data point is modeled to the map node by calculating the positions of 1<sup>st</sup> BMU and 2<sup>nd</sup> BMU. If these are located next to each other, the topology is preserved, and the TE is said to be zero for individual input. Similarly, summing up the errors for every input and calculating the data points as average are considered TE for the map as follows:

$$\text{TE}(\mathbf{M}) = \frac{1}{\mathbf{N}} \sum_{i=n}^n \mathbf{t}(\mathbf{x}_i), \quad (7)$$

where  $\mathbf{t}(\mathbf{x}_i)$  is a piecewise function; it is 0 if 1<sup>st</sup> BMU and 2<sup>nd</sup> BMU are neighbors or 1 otherwise. TE is evaluated to quantify the topology preservation by evaluating local discontinuities in the output map. Mostly, a tradeoff is realized between quality measures when increasing the projection quality and seems to decrease when some information is lost during this process.

### 3.2.2. Quantum Phenomenon

(1) *Quantum Computing-Overview.* The QC is based upon physics' natural laws and claims to solve many (sub) atomic level problems that are inflexible for old style computers. Quantum parallelism is a distinctive feature established on superposition and entanglement and offers exponential acceleration of computation over conventional computers, especially for cryptosystems, making them acutely fast [30]. A quantum computer taking " $n$ " qubits that can be stored and execute  $2^n$  imaginable combinations of qubits simultaneously by joining them in an uncommon fashion recognized as superposition and defined as follows:

$$|\Psi\rangle = \alpha_0|0 \dots 00\rangle + \alpha_1|0 \dots 01\rangle + \dots + \alpha_{2^n-1}|1 \dots 11\rangle, \quad (8)$$

where  $\alpha_i \in \mathbb{C}$  complex numbers known as probability amplitudes of qubits and  $\sum |\alpha_i|^2 = 1$ .

In quantum information processing, electrons or photons in a coherent state is encoded with some required information (known as qubits) and pass to another qubit via a quantum bus. The passed information is accessible to many qubits in a system, accelerating the speed of computation, unlike classical computation [31, 32]. Trapped ion architecture, QC using superconducting qubits, and QC with nitrogen-vacancy center in diamonds are few of the hardware architectures considered for thorough research in well-equipped labs [33].

Like classical computation, quantum computation is carried out with the help of quantum gates. The information that has been obtained from quantum gates can be reversed. The representation of a single qubit quantum system is in the

form of a Bloch (shown by Figure 3). Each quantum gate is represented by a matrix containing complex coefficients and can be applied on the qubit (state vector in Bloch sphere) to change state as a vector. A qubit is a state vector in two-dimensional Hilbert space. This vector can have any direction from the sphere's center to its periphery, i.e., it can connect to any point on the sphere's surface. The poles indicate the ground and excited states, and anywhere between these points is the superposed state of the qubit. Different qubit transitions indicate rotations about the axes changing the state of that qubit [34, 35]. Moreover, these rotations occur as a result of the quantum gate(s), which act on that qubit (see Figure 3).

(2) *Two-Level Spin System*. As described earlier, logical content prioritization is achieved through the DL agent to know about the requested content's priorities. Once the priority is known, the particle is encoded according to the relevant content priority. By taking into account the high and medium priorities, the contents can be stored physically in QMM by dividing it into two groups for which SGE is exploited. The information is encoded through the spin  $\pm 1/2$  particles;  $+1/2$  upward spin and  $-1/2$  downward spin. The highly prioritized data are coded by spin-up particles, whereas through the spin-down particles, the likewise and medium prioritized data are encoded.

Due to the spinning environment, the electron has a magnetic field. The magnetic field can be canceled by another electron having an opposite spin in an atom. In an atom, the electrons are either paired or unpaired. To decide the spinning effect, the unpaired electrons leave the orbitals. Electrons are accrued like a beam, divided into two illustrious beams of equivalent power even though passing through a nonuniform magnetic field. Therefore, a massive tendency is shown in Figure 4.

A quantum organization is indicated by its state vector. But at times, the system is said to be in mixed state having statistical ensemble of various state vectors. Such a system has equal probabilities or chances to designate in either pure state. The pure state is basically a quantum state useful in the quantum system and determines the statistical behavior of the measurement. At the beginning of the TLSQP, all electrons are located in a mixed state since the states are indefinite. Due to the half-half chances of existence in any of the pure states ( $|0\rangle$  and  $|1\rangle$ ), the particles have a mixed state. Density matrices are used to represent the statistical state of the quantum system or a particle. The chances for result can be calculated from the density matrix for the system. The density matrices for states (i.e., mixed and pure) represented that the particles are initially in the mixed state when accrued like a single beam.

$$\rho = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}. \quad (9)$$

A nonuniform magnetic ground is formed by employing two magnets in a perpendicular way along the  $z$ -axis. When

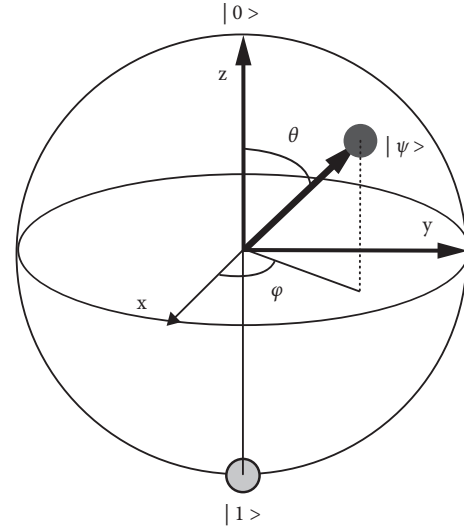


FIGURE 3: Representation of a qubit in a 2-dimensional Bloch sphere.

a beam passes through the magnetic field, electrons are bent alongside the axis comparative to the  $z$ -component. After passing through a magnetic field, some of the particles are in the Eigen state  $|Z+\rangle$  of the  $S_z$  operator. The matrix for this state vector is given as follows:

$$\rho_{Z+} = |Z+\rangle\langle Z+| = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}. \quad (10)$$

The trace of this density matrix's square is 1, which clearly shows that it is a pure state. The state of the rest of the particles after passing through the magnetic field becomes 1 ( $|Z+\rangle$ ). The matrix for this state is given as follows.

$$\rho_{Z-} = |Z-\rangle\langle Z-| = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}. \quad (11)$$

Again, this density matrix's trace is 1 and shows that it is a pure state. In contrast, if considered these two separated beams together, we again get the mixed state consisting of an equal mixture of particles in Eigen states  $|Z+\rangle$  and  $|Z-\rangle$ , as follows:

$$\rho = \frac{1}{2}|Z+\rangle\langle Z+| + \frac{1}{2}|Z-\rangle\langle Z-| = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}. \quad (12)$$

The trace of the square of this density matrix is  $1/2$  which is less than 1, showing a mixed state. Before passing via a magnetic field, the electrons' spin can point in any direction of the space, being equally probable, so there is no state vector but for pure states.

The mixed beams can be stored in QMM distinctly. The  $|Z+\rangle$  state represents Repository 1 electrons devising pure state  $|0\rangle$  wherever (high) prioritized contents are stored. The  $|Z-\rangle$  state signifies Repository 2 electrons taking pure state  $|1\rangle$ , and the intermediate prioritized



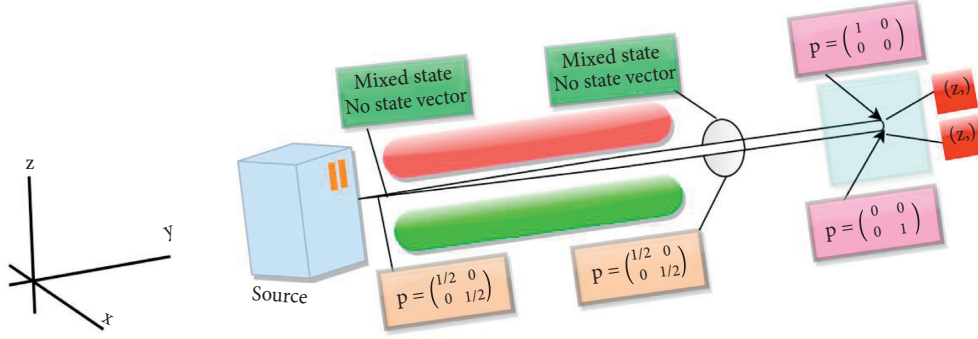


FIGURE 4: Electron beam splitting into two while passing through a magnetic field [23].

contents are stored. Therefore, electrons containing certain contents' information are categorized based on the established priorities by using TLSQP. The stated repositories help in storing the modified contents in every interval of time.

The ion-trap architecture of QMM is useful and effective for this specific scenario. The quantum data can be stored by qubits using atomic ions. The qubits (atomic ions) are trapped and designed by groupings of static and oscillating electric fields [33, 36, 37]. In what way, these quantum data are stored in these repositories which are beyond the scope of this research. The quantum information can be managed (processed or transferred) through the ions' cooperative quantized motion and is also recognized as quantum parallelism. The respective parallelism leads to an increase in the processing time as compared to the classical architectures. It has long been known that classical physics principles do not allow for causally efficacious understanding; yet, the intrinsic indeterminism and characteristic duality of quantum physics is that it contains give fertile ground for comprehension through physical modeling.

Measuring probability for spin-up and spin-down particles is an important aspect that will help determine the category of data encoded in a particular spin-type particle. So, to interpret the idea, IBM's QC simulator is exploited to yield some meaningful results. The resulting probability  $p$  for a particle to emerge as a spin-up particle can be found out by  $\cos^2(\theta/2)$ , and for a particle to emerge as spin-down can be found out by  $\sin^2(\theta/2)$ , where  $\theta$  is the angle of rotation along  $Z$ -axis. Different angles of rotation will yield different likelihoods for spin-up and spin-down elements. The formula to find the probabilities of spin-up and spin-down elements is explained as the trace of density matrix and projection operator on that pure state-directed to some angle  $\theta$  as follows:

$$\mathbf{p} = \text{Tr} [\rho \mathbf{P}_{\hat{n}}], \quad (13)$$

where  $\rho$  is the density matrix of pure states already described above and  $\mathbf{P}_{\hat{n}}$  is the projection operator on a pure state which is directed to some  $\hat{n}$  so that  $\hat{n}_\theta = (\cos\theta, 0, \sin\theta)$ . Hence,

$$\begin{aligned} \mathbf{P}_{\hat{n}} &= \frac{1}{2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 0 & \sin\theta \\ \sin\theta & 0 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} \cos\theta & 0 \\ 0 & -\cos\theta \end{bmatrix} \\ &= \begin{bmatrix} \cos^2\frac{\theta}{2} & \frac{1}{2}\sin\theta \\ \frac{1}{2}\sin\theta & \sin^2\frac{\theta}{2} \end{bmatrix}. \end{aligned} \quad (14)$$

So, the trace of the product of  $\rho$  and  $\mathbf{P}_{\hat{n}}$  comes out to be  $\cos^2(\theta/2)$  for  $|Z+\rangle$  particles and  $\sin^2(\theta/2)$  for  $|Z-\rangle$  particles, respectively.

Quantum computers have stimulated the rotation produced by the magnetic field in the SGE by applying quantum gates. For instance, a **T** gate is used to produce rotation at  $\theta = \pi/4$ . This gate rotates the state of the qubit in the superposed form by angle  $\pi/4$  along the  $Z$ -axis. So, it is necessary to apply the Hadamard gate (**H**) before applying the **T** gate, as the **H** gate helps create superposition. Matrix representation of a Hadamard gate is shown as follows:

$$\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \quad (15)$$

It converts the  $|0\rangle$  basis state of the qubit to  $((|0\rangle + |1\rangle)/\sqrt{2})$  form (also known as  $|+\rangle$ ), and  $|1\rangle$  basis state to  $((|0\rangle - |1\rangle)/\sqrt{2})$  form (also recognized as  $|-\rangle$ ). Therefore, a superposition is created as there is an equal probability to be either 0 or 1. It produces two rotations simultaneously:  $\pi$  at the  $z$ -axis and  $\pi/2$  at the  $y$ -axis and is shown by Figure 5(a).

After creating superposition, the **T** gate is applied to rotate the superposed qubit at  $\pi/4$  along the  $z$ -axis. This is a single qubit gate (from the family of phase shift gates), which does not change the probability of the  $|Z+\rangle$  and  $|Z-\rangle$  somewhat changing the phase of the qubit's state. This gate acts on the  $|1\rangle$  base state, whereas exiting the  $|0\rangle$  base state remains unaffected. So,  $|+\rangle$  will be converted to  $((|0\rangle + i\pi/4|1\rangle)/\sqrt{2})$  and  $|-\rangle$  will be converted to  $((|0\rangle - i\pi/4|1\rangle)/\sqrt{2})$  because these are mixed states. Matrix representation of **T** gate is described as follows:

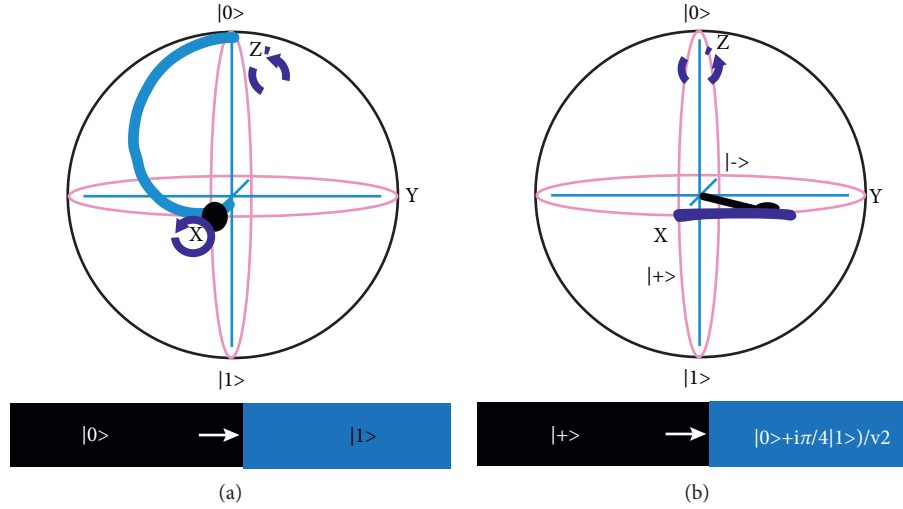


FIGURE 5: The gates are represented through a Bloch sphere. (a) Hadamard (H) gate representation in a Bloch sphere, (b) The rotation is produced by T gate and shown in a Bloch sphere.

$$\begin{bmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{bmatrix}. \quad (16)$$

The rotation produced through the T gate and is shown in Figure 5(b).

As soon as the T gate is applied, another H gate is again useful to maintain qubit's superposition. Alternatively, it would have lost its quantum state and collapsed into a classical one that is of course 0 or 1 depending upon the qubit chosen. The equation which satisfies this circuit is given as follows:

$$\frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} + \frac{1}{2}e^{i\pi/4} \\ \frac{1}{2} - \frac{1}{2}e^{i\pi/4} \end{bmatrix}. \quad (17)$$

When more than one gate is applied on a qubit in a serially wired circuit, dot product (usual matrix multiplication) is carried out for all the gates, resulting in a combined gate acting on that qubit. As mentioned in equation (17), H, T, and H gates have been combined by dot product and applied on  $|0\rangle$ . The resultant matrix shows the probability amplitudes of spin-up and spin-down, as complex numbers, just before the measurement. It can also be written as follows:

$$\left(\frac{1}{2} + \frac{1}{2}e^{i\pi/4}\right)|0\rangle + \left(\frac{1}{2} - \frac{1}{2}e^{i\pi/4}\right)|1\rangle. \quad (18)$$

The probability amplitude  $\alpha^2$  of  $|0\rangle$  is  $|(1/2) + (1/2)e^{i\pi/4}|^2$ , equal to 0.853534, approximately 85%, and that of  $|1\rangle$  is the leftover probability which is of course 0.146466 (approximately 15%). Hence,  $|0\rangle$  basis state has a greater probability, so classically, a 0 is obtained by measuring the state if the T gate is applied. Similarly, some other appropriate sequence of H gates and phase shift gates can also be applied in order to produce a distinct rotation and obtain

different probabilities of spin-up and spin-down particles. It depends upon which type of particle is needed to encode the data to be stored in the relevant repository.

A quantum computer can help to determine these complex probability amplitudes in terms of real numbers. It can then be classically interpreted and ultimately helping in encoding data.

**3.3. Flowchart and Algorithm (Solution).** The flowchart of the proposed framework is represented in Figure 6. An algorithm is described (and also shown by Algorithm 1) as follows.

The algorithm comprises three functions: (1) cache\_synchronization [], (2) cache\_update [], and (3) Serve\_UE\_Phase []. The cache\_synchronization [] function is used for cache synchronization and cooperation. It has parameters  $t$ , and  $S$ :  $t$  is the time interval after which cache information is shared, while  $S$  is a set of regional users for a particular  $t$ . It returns the regional user set for a particular time interval by considering the time interval, set of user requests, the workload on the edge node, and distance  $d$  of UE from the edge node. The information of  $R$ ,  $d$ , and  $w_E$  at a specific time, quantum  $t$  is shared in Step 1.  $R$  is used for the set of user requests  $\{r_1, r_2, \dots, r_k, \dots, r_n\}$ ,  $d$  is the distance of UE from edge node receiving a request, and  $w_E$  is the workload on a particular edge node. Step 2 will return a list of  $S$  for the particular  $t$ . In Step 3, assigned edge node  $E_A$  to a particular user set,  $S$ . Step 4 is used as a counter for  $t$  and Step 5 will repeat Step 1.

The cache\_update [] function is used to update the caching contents. It consumes (as input) the list  $L$  of contents with extracted features and produces (as output) the prioritized caching content to be placed in F-AP. Step 1 selects the appropriate map size: horizontal  $x$  and vertical  $y$  dimensions of the map. Step 2 defines the color intensity of map nodes to depict classes for low, medium, and high priority contents. Step 3 is used to run the SOMs algorithm

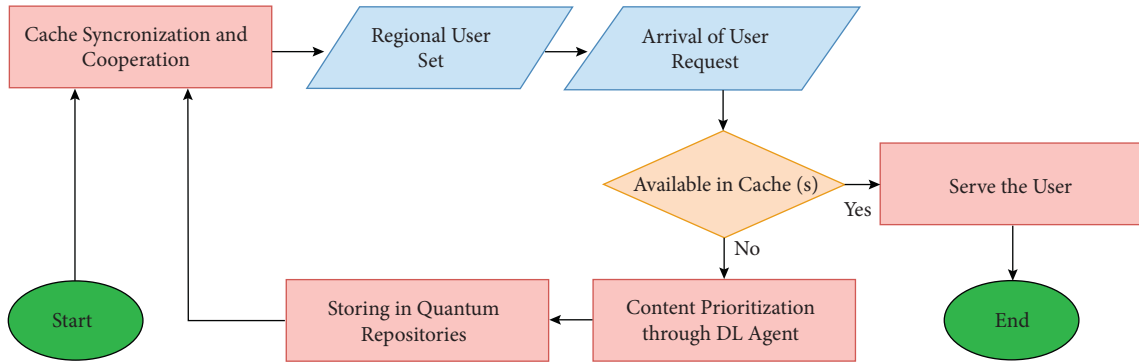


FIGURE 6: The flowchart of the proposed framework.

after initializing weights  $W_{ij}$ . The mapping of  $n_{DB}$  and  $n_{MB}$  on  $L$  to get  $C_p$  in the Step 4. The map nodes with dark blue  $n_{DB}$  and medium blue  $n_{MB}$  color showing high priority and medium priority content as  $L$ .  $C_p$  are the prioritized contents to be placed in cache.

Serve\_UE\_Phase [] function is used to serve a user request  $r_k$ . It takes regional user set and assigned edge node as input and activate to serve for an incoming request. If there is a cache hit in Step 1, it will serve  $r_k$ ; otherwise, it will first update the cache and then serve  $r_k$ .

#### 4. Content Prioritization Results through Deep Learning

**4.1. Self-Organizing Maps.** Multimedia content needs to be prioritized with considerable views, likes, dislikes, and comments. SOMs algorithm learns from the given dataset and displays on the map by the grid of nodes. The degree of the relationship between data points is shown through the color intensity. As a proof of concept, a tool is implemented using Python programming language and is exploited through a Jupyter notebook in Google Colab. As mentioned earlier, the four input vectors, i.e., views, likes, dislikes, and comment\_count, are selected utilizing the dataset to simulate. The feature scaling is achieved through the MinMax scaler. To train the SOMs algorithm, the tuning parameters with their values are simulated through the tool and shown in Table 1.

To select an appropriate lattice size, different lattice sizes are tested by the hit and trial method to validate the formula. The experiment shows that the batch training yields many exact results; however, it is a bit slower than the random training. The recorded data are shown in Table 2. A histogram also represents the recorded data in Figure 7. On  $x$ -axis, different lattice sizes (i.e.,  $10 \times 10$ ,  $15 \times 15$ ,  $20 \times 20$ , and  $32 \times 32$ ) with different number of iterations (i.e., 250,000, 500,000, 750,000, and 1,000,000) are shown by different colors. On the  $y$ -axis, the error values are displayed (given in Table 2). Evaluating the data carefully proves that the error values are recorded minimum on the lattice size  $32 \times 32$  with 750,000 iterations; therefore, this lattice size is considered appropriate.

The map's outputs of different iterations for lattice size  $32 \times 32$  are shown in Figure 8. The color scale for iterations 250,000, 500,000, 750,000, and 1,000,000 are shown in Figures 8(a)–8(d), respectively. Nodes with color (values)

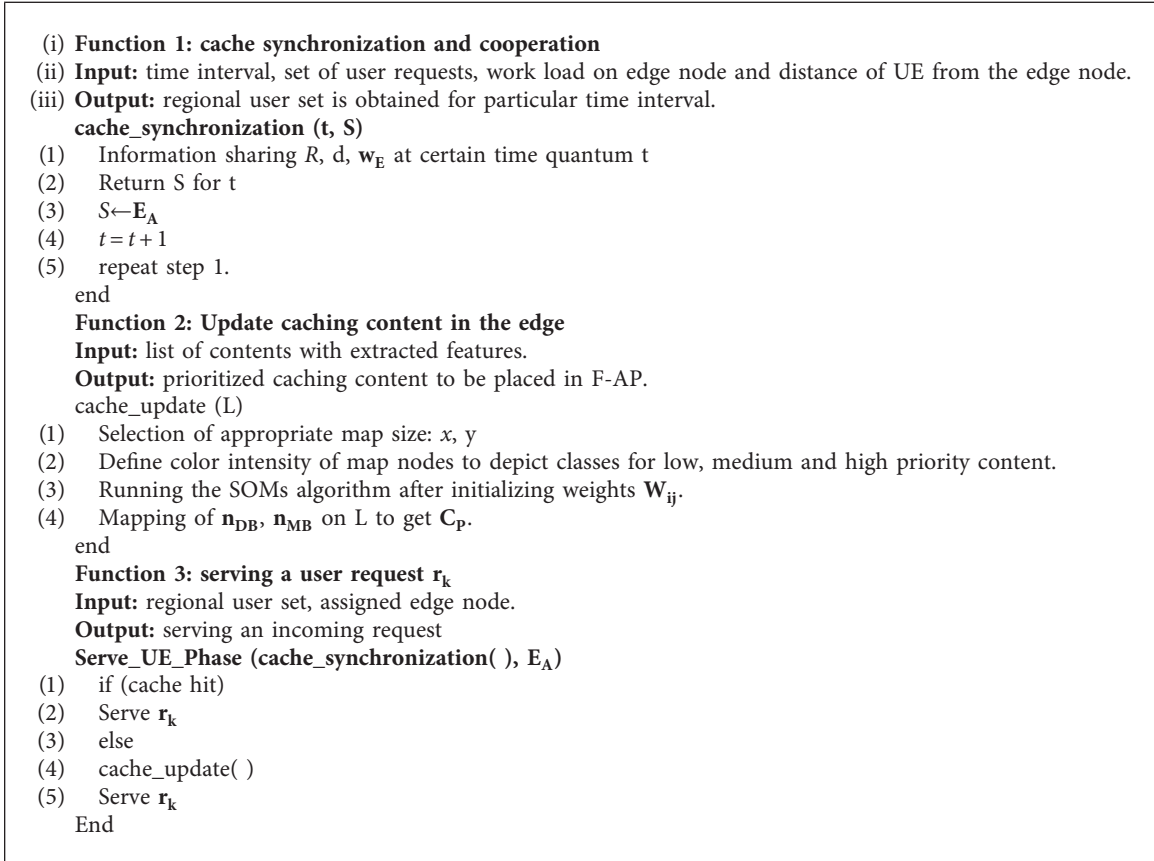
range from 0.8 to 1.0 (dark blue) which represent the group of data having high priority contents (maximum number of views, likes, dislikes, and comments). The medium priority contents are represented with light bluish color nodes ranging from 0.4 to 0.8. The medium priority contents follow the high priority contents. The remaining nodes with color values below 0.4 are considered the least priority contents and must not be deliberated in the caches.

After the SOMs algorithm's training through different lattice sizes, the QE needs to be extracted to check the validity of the data. According to the data described in Table 2, the best map among all is  $32 \times 32$  lattice-sized map with 750,000 iterations. The QE value is recorded even less than 0.000024 as shown by the graph in Figure 9(a). It is not reduced any further after 0.000024 QE value. The TE is plotted by Figure 9(b) to determine how well the topology of the map is preserved at 750,000 iterations. The TE value at this point is logged as 0.092. Although the TE is recorded a little bit higher but its value, together with QE value (i.e., 0.0000235) plotted the most appropriate map after 750,000 iterations. By taking into account the curves shown in Figures 9(a) and 9(b), it can be analyzed that the map is trained efficiently and delivers precise results.

After selecting the most appropriate lattice map ( $32 \times 32$ ) in 750,000 iterations, the data points located below the dark blue region are mapped onto the videos' data frame. Similarly, the light-blue color region is also mapped to get medium prioritized content. Table 3 is used to describe the mapping data of high priority contents from one of the nodes from the dark blue region (27, 2). Also, it shows twelve highly trending videos (in rows from 0 to 11) with respect to views, likes, dislikes, and comment\_count (in columns). These videos are considered as a high priority for trending according to the input parameters provided in the dataset.

The rest of the nodes from bluish-white to white are located in the lighter region and can be ignored because this region contains the least priority content.

**4.2. Quantum Self-Organizing Maps.** Quantum-SOM (QuSOM) has a different learning method than SOM. The number of presynaptic neurons corresponds to the number of neurons in both layers and interconnections between them when designing the QuSOM layout, which comprises



ALGORITHM 1: Cache content management.

TABLE 1: Tuning parameters and their values.

Tuning parameters with symbols	Values
$x$ -dimension of the lattice ( $x$ )	32
$y$ -dimension of the lattice ( $y$ )	32
Learning rate ( $L_t$ )	0.1
Initial spread value ( $\sigma_0$ )	1.0

TABLE 2: The errors' values with respect to different lattice sizes and the number of iterations.

Dimensions of lattice	No. of iterations			
	250,000	500,000	750,000	1,000,000
$10 \times 10$	0.0001555	0.0001563	0.0001584	0.0001643
$15 \times 15$	0.0001218	0.0001110	0.0000970	0.0000890
$20 \times 20$	0.0000847	0.0000727	0.0000594	0.0000460
$32 \times 32$	0.0000355	0.0000353	0.0000235	0.0000241

the number of all model parameters and the set of potential data classifications. There is only one connection between a neuron in the input layer and a neuron in the output layer. QuSOM attracts all vectors of  $v$ , ( $v(i) \in i, i = 1, 2, \dots, N$ ), just once. The competitive and weight update is accomplished through a series of procedures, which is a parallel processing capability. As a result, in QuSOM, the traditional repetitive learning procedure is modified to learn only once. Algorithm 2 of the QuSOM is as follows:

The QuSOM, in QC, may shift research directions in the artificial neural network (ANN) field depending on the computation environment and application property [38]. The parallelism aspect of the QuSOM is its most intriguing feature. A quantum mechanics computer can exist in a state of superposition and perform several operations simultaneously. A QuSOM gate array is depicted in Figure 10 as a schematic. The register's initial state is on the left, and time moves from left to right. The  $W^s$  gate is a weight operator at  $s$ ;  $D^s$  gate is a distance operator at  $s$ ;  $d^s(i, j_{\min})$  is a Grover searching oracle;  $W^{s+1}$  is a winner weight updating operator at  $t$ ;  $U$  is a weight transformation operator at  $s, u = QW^{t+1}$ ;  $\vartheta$  is an observable extracted information from register, according to the above summarized QuSOM algorithm. The operations of these transformation and operation matrices are used to create QuSOM. The vectors are only entered into the map once, and the output (weight) should converge if the sequence is repeated.

In the traditional meaning of computation, putting all parameters in inputs as neurons may be unfeasible, and QuSOM operation will be time consuming due to parallelism.  $N = 2$  input vectors with  $M = 4$  total input items and  $P = 2$  prototypes, for example, and the number of neurons in both input and output layers should be  $2 \times 4 \times 2 = 16$ . This figure is four times that of SOM. Fortunately, this is not an issue in quantum computing. Quantum theory's peculiar properties can be used to express information with a neuron

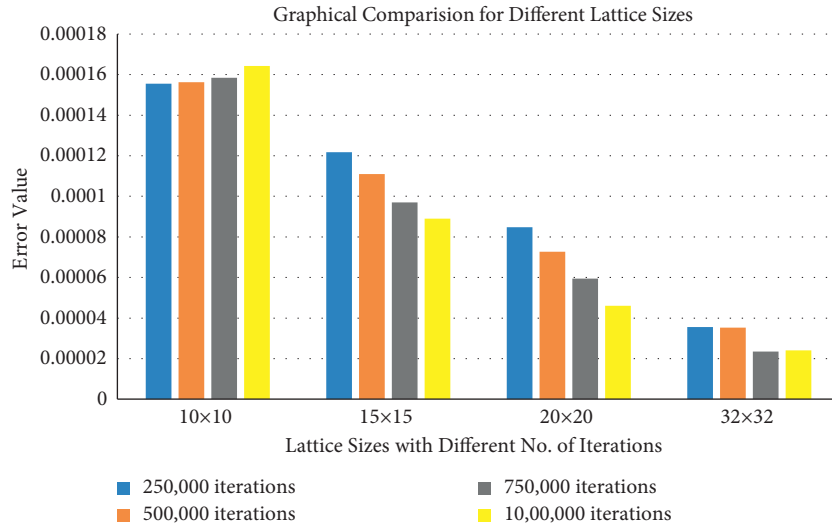


FIGURE 7: Comparison of different lattice sizes with the errors values by selecting different number of iterations.

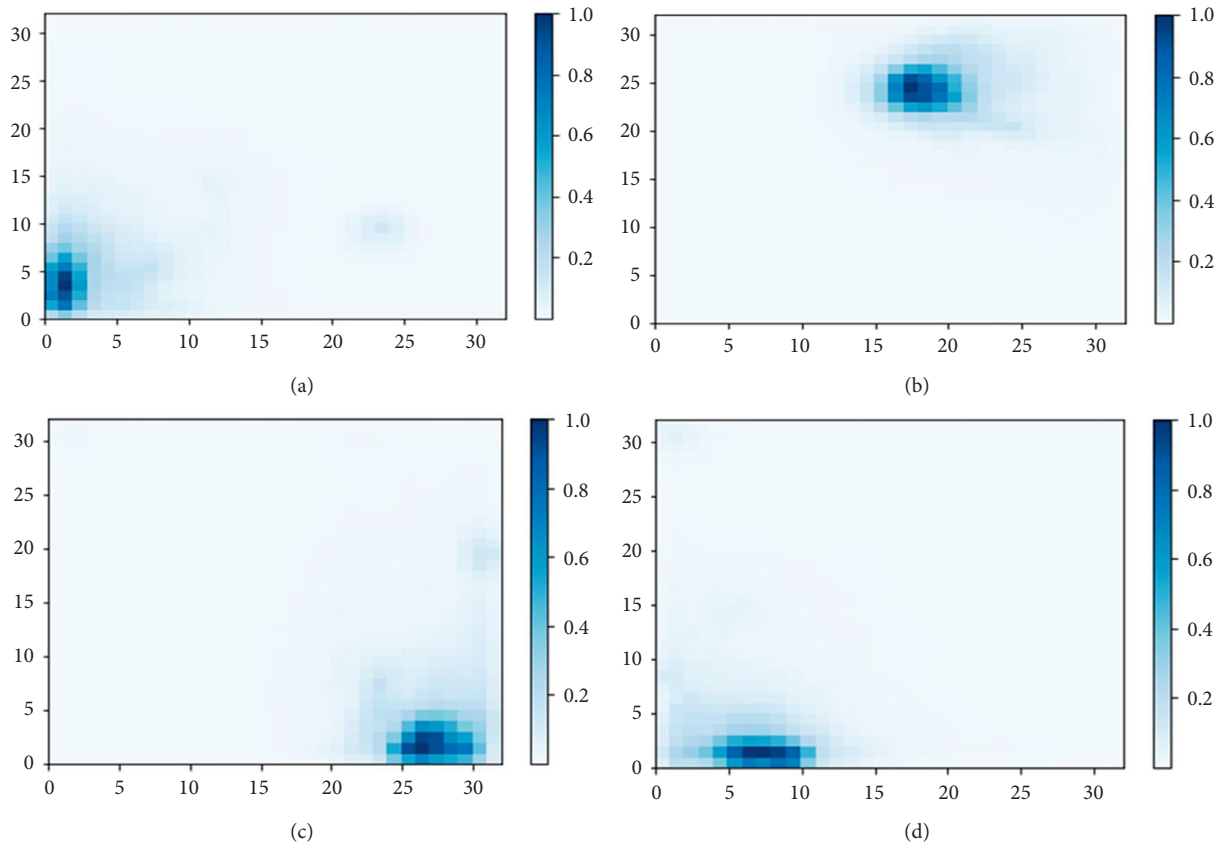


FIGURE 8: The color scales for different iterations at lattice size  $32 \times 32$  are shown. (a) Output map at 250,000 iterations. (b) Output map at 500,000 iterations. (c) Output map at 750,000 iterations. (d) Output map at 1,000,000 iterations.

number of exponential capacity. The number of neurons is exponentially decreased to  $\log_2$  for the input signals,  $v(i, k)$ ,  $i = 1, \dots, Z$ ,  $k = 1, \dots, Y$ ,  $j = 1, \dots, P$ , by adopting quantum representation ( $M \times N \times P$ ). QuSOM requires only 4 qubits in the example above. The input data from YouTube streaming may be more than 4 vectors with 40960 elements in some edge caching systems. The SOM configuration was used with

4 neurons in the input layer and  $32 \times 32 = 1024$  neurons in the output layer. The network might be configured with 27 quantum input/output neurons, or qubits, representing roughly 41943040 SOM neurons using QuSOM. So, it can be determined that the gain difference in terms of computation and time consumption between the deep learning method based on quantum computing and conventional method as



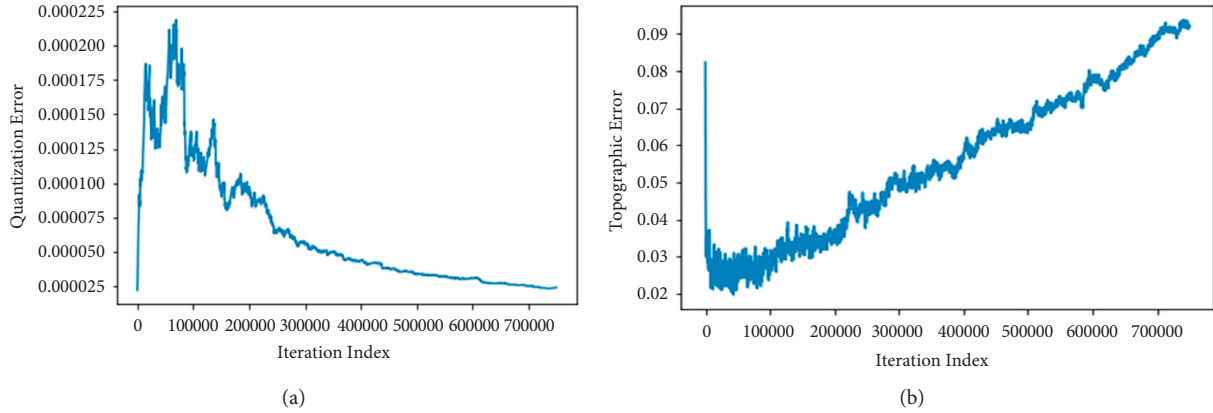


FIGURE 9: (a) Quantization error to check the validity of the data and (b) topographic error to determine how well the topology of the map is preserved at 750,000 iterations.

TABLE 3: The data of the Mapping Dark Blue Node (27, 2).

	Views	Likes	Dislikes	Comment_count
0	167997997.0	4281819.0	276626.0	453206.0
1	173478072.0	4360121.0	283961.0	460299.0
2	179045286.0	4437175.0	291098.0	466470.0
3	184446490.0	4512326.0	298157.0	473039.0
4	190950401.0	4594931.0	305435.0	479917.0
5	196222618.0	4656929.0	311042.0	485797.0
6	200820941.0	4714942.0	316129.0	491005.0
7	205643016.0	4776680.0	321493.0	496211.0
8	210338856.0	4836448.0	326902.0	501722.0
9	217750076.0	4934188.0	335462.0	509799.0
10	220490543.0	4962403.0	338105.0	512337.0
11	225211923.0	5023450.0	343541.0	517232.0

- (1) **Input vector**  $V = (v_1, v_2, \dots, v_n)$ , i.e., QuSOMs learn all of information simultaneously.
- (2) One operation of **competitive** and **updating**, for  $s$  iterations
- (3)  $W^s = U$  (in first operation,  $W^1 = W^0$ );
- (4)  $D^s = \|V - W^s\|$
- (5)  $d^s(i, j_{\min}) = \min(d^s(i, 1), d^s(i, 2), \dots, d^s(i, P))$ ;  $i = 1, 2, \dots, Z$ ;
- (6)  $w^{s+1}(i, j_{\min}) = w^s(i, j_{\min}) + \eta(s)[x(i, j_{\min}) - w^s(i, j_{\min})]$ , if  $j = j_{\min}$ ,
- (7)  $w^{s+1}(i, j) = w^s(i, j)$
- (8) If  $W^{s+1} - W^s > \epsilon$  go to 9, otherwise go to step 10
- (9)  $V = Q W^{s+1}$  go to 3
- (10) **Store**  $W^{s+1}$  and **stop**

ALGORITHM 2: Learning through quantum self-organizing maps.

the measures of conventional computing are almost four times the quantum computing that clearly shows extensive use of resources in conventional computing.

#### 4.3. Simulation Results for Quantum Phenomenon in Caching.

For the experiments, a cloud-based QC system (from IBM Quantum Experience) is used. The IBMQ\_QASM\_Simulator is basically a simulator backend, allows sampling circuits with a 32 qubits processor. The circuit to produce rotation of the particles is shown in Figure 11. It is a serially wired circuit comprising three gates (two

Hadamard and one  $T$  gate). The respective circuit is used to act on qubit  $|0\rangle$  and is known as a standardized measurement operator along the  $z$ -axis. The primary (first) wire, labeled as  $q[0]$ , is a quantum wire representing the passage of time. It is not considered a physical wire. The gates are applied in unit time. The second wire, labeled as  $c1$ , is a classical wire, and the output from the quantum computer is determined once the measurement is applied. The vertical arrow from the measurement operator shows that the information is now retrieved from the quantum regime to the classical regime.



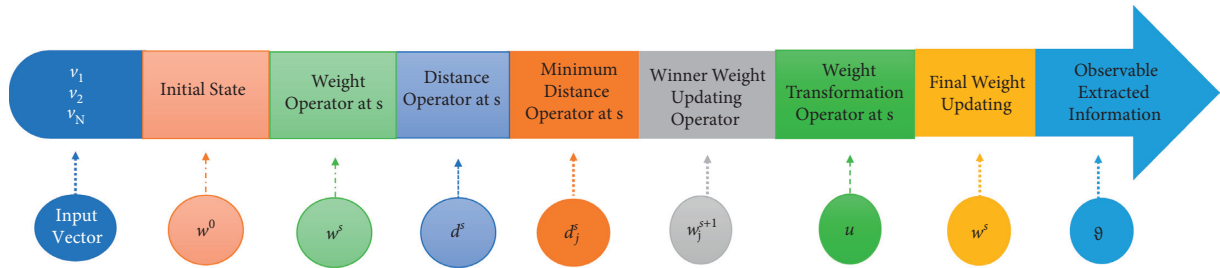


FIGURE 10: The general QuSOM structure contains gate array.

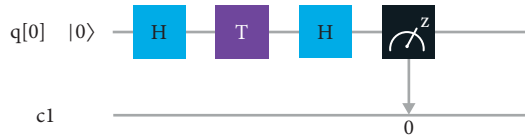


FIGURE 11: The circuit to produce rotation of the particles is shown from IBM QASM Simulator.

To evaluate circuit on the simulator, a parameter *number of shots* is needed to set before execution. The *number of shots* is simply a parameter having a value that determines the number of iterations, representing the number of times a quantum circuit executes. With the increase in the *number of shots*, the probability values of spin-up and spin-down are improved. The resultant probabilities are demonstrated by Figure 12 with different numbers of shots (i.e., 1024, 4096, and 8192 shots) and actual measurements. The horizontal axis of the histogram represents the computational basis states 0 and 1. The vertical axis represents the probability of observing that basis state. The histogram in Figure 12 also represents the exact probability measurement of the basis states.

As clearly depicted from Figure 12, there is a slight difference between the actual and simulated results. The first and last simulated results are realized by 1024 and 8192 shots, respectively. The probability measurement difference is reduced as compared with the actual result. The probability measurements are close to the actual result by 8192 shots. The results are concluded (by comparing actual and simulated) in Table 4.

All the prioritized contents are grouped in the form of color clusters. The color intensity of the nodes in the feature map makes it easier to prioritize the content. The nodes with less intense color illustrated the low prioritized content. The selection of highly prioritized and medium prioritized contents is achieved conveniently through the SOMs algorithm. The algorithm took less time and computational overhead than other DL algorithms. The use of TLSQP facilitates the overall management of most requested content within the edge for providing an instantaneous content delivery response. The concept of storing content in QMM by employing this quantum phenomenon is entirely unconventional and challenging at the same time. Nevertheless, its advantage overshadows other conventional approaches of the classical regime in terms of storage capacity and processing speed due to its unusual properties, i.e., quantum parallelism.

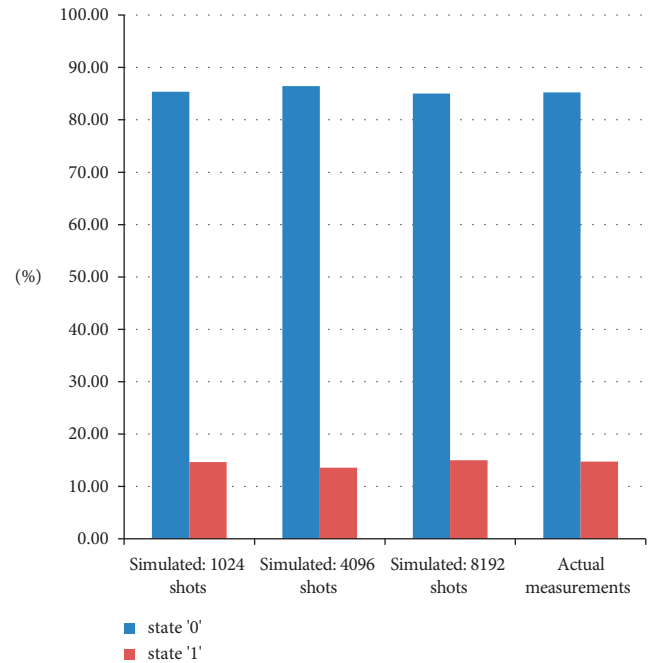


FIGURE 12: The simulated resultant probabilities with different number of shots, i.e., 1024 shots; 4096 shots; and 8192 shots; and the actual probability measurements are shown.

## 5. Discussion

The framework is basically a merger of a DL agent deployed at the network edge and a QMM. Firstly, the DL agent prioritizes caching contents via SOMs algorithm, and secondly, the prioritized contents are stored in QMM using TLSQP. The study of QuSOM follows the development tendency of ANN. The adaptation of ANN in the parallel computing environment will be interesting for both ANN and QC, especially for the simulation of human learning and memorizing features by using more powerful computing tools. In Kohonen's SOM, the learning and weight updating are organized in the same sequence. This sequence is like the human's repeated learning manner. In QuSOM, due to its once learning property, the weight updating is managed separately with learning and updating. This manner may appear more similar to human's once learning way. QuSOM has the same convergence property as Kohonen's SOM, but its time and space complexities are more simplified. To verify the valuation and efficiency of the algorithm, in this study, we have compared the gain difference in terms of

TABLE 4: Comparison of actual and simulated probability measurements.

Basic states	Probability measurements			
	Actual (%)	Simulated (1024 shots)	Simulated (4096 shots) (%)	S (%) imulated (8192 shots)
0	85.35534	86.426	85.01	85.242
1	14.64466	13.574	14.99	14.758

TABLE 5: Comparison with previous research.

Research	Deployed algorithm	Performance measurements	Results	
[39]	Multiagent deep reinforcement learning (MADRL)	Caching reward	21%	
		Cache hit rate	Highest	
		Traffic load	43%	
		Caching reward	56%	
		Cache hit rate	Higher	
		Traffic load	45%	
	Deep reinforcement learning (DRL)	Caching reward	43%	
		Cache hit rate	Lower	
		Traffic load	36%	
		Caching reward	34%	
		Cache hit rate	Lowest	
		Traffic load	12%	
[40]	Personalized edge caching system (PECS)	Deep packet inspection	Top-down analysis (network level) and bottom-up analysis (user level)	
[41]	One-dimensional convolutional neural network (ODCNN) Self-Organizing Map (SOM)	Accuracy rate	99.8%	
[42]	Support vector machine	Accuracy	0.984	
		Precision	0.984	
		Recall	0.983	
		F1 score	0.981	
		Accuracy	0.983	
		Precision	0.982	
	Logistic regression	Recall	0.983	
		F1 score	0.983	
		Accuracy	0.984	
		Precision	0.983	
		Recall	0.984	
		F1 score	0.984	
K-nearest neighbors	Accuracy	0.870		
	Precision	0.969		
	Recall	0.973		
	F1 score	0.919		
	Proposed	Self oranizing map (SOM)	Quantization error	0.000024
			Topographic error	0.092
QE + TE			0.0000235	
Two-level spin quantum phenomenon (TLSQP)	Basic states	0	OV = 85.4% PV = 85.2%	
		1	OV = 14.6% PV = 14.8%	

computation and time consumption between the deep learning method based on quantum computing and conventional method which can be summarized as the measures of conventional computing are almost four times the quantum computing.

To verify the algorithm, we conduct extensive experiments to demonstrate that the algorithm improves the generalizability of the conventional SOM through optimization and is robust to the choice of hyperparameters, as listed in Table 5.

A hybrid approach was used for this work: one (i.e., SOM) for classifying the videos dataset and a second (i.e., TLSQP) for the storage of prioritized content. It can be inferred that the proposed framework DL-QC is deployed for edge caching in sensor network traffic to improve the prioritization and storage processes by exploiting the capabilities of DL and QC. Dataset has been selected that incorporates multimedia content that has been infrequently used in the past for other studies. The dataset contains four features and 40960 samples. Google Colab and IBM's

Quantum Experience are utilized in this work with high certainty because of their capabilities of creating legitimate outcomes that mirror certain domains of intelligence and quantum. The gathered information has been recorded for prioritization levels and will notify the prioritized cases to the QMM storage through TLSQP, whereas the most minor prioritized cases will be removed with a higher accuracy rate. The DL algorithm SOM is precisely applied to the identified dataset for prioritization in a  $32 \times 32$  lattice size. The selected DL classifiers accomplished the particular task with accuracy and precision, as discussed above. It will help characterize the high priority and medium priority network traffic to ensure the optimized caching services in the edge computing environment, and TLSQP ensures maximum data storage in QMM. Due to the research scope, some common types of multimedia content parameters have been selected, but in the future, more categories of innovative content and features in the DL-QC environment can be incorporated to better understand and cope with the identified issues. Furthermore, innovative algorithms can also be designed, or existing ones can be modified to prioritize and storage than already discussed to improve efficiency and accuracy. Moreover, the QuSOM can be replicated on conventional computers as well as quantum computers provided that the availability of resources to understand the results better.

## 6. Conclusion

The caching content's storage is mainly the primary source of immediate delivery responses. This research work has presented an intelligent DLAQC framework for updating the EC content in F-RANs. The caching content is logically prioritized through an intelligent DL agent in the network edge using the SOMs algorithm. The caching content is physically stored in QMM, exploiting the TLSQP phenomenon to update the caches and provide ample content storage for immediate delivery response against the unpredicted amount of static and dynamic user requests. The framework is evaluated using multimedia content and provides effective outcomes, especially by reducing computation overhead and time. The purpose is to form clusters to separate high, medium, and low-prioritized contents in an unsupervised manner. SOMs algorithm is staunchly suitable to pick up contents in colored cluster form without reducing the dimensionality of the feature space.

While the experiments have been conducted for multimedia content only, other contents can be considered, especially in IoT-based scenarios where unpredictable amounts of static and dynamic requests are generated day by day. EC is capable of handling each request immediately; it is still challenging and can be considered to explore further. Besides, innovative algorithms can also be designed, or existing can be modified to prioritize and storage than already discussed to get better efficiency and accuracy. Moreover, the QuSOM can be replicated on conventional computers as well as quantum computers provided the availability of resources to better understand the results.

## Data Availability

A publicly available dataset is used for this study (Trending YouTube Video Statistics).

## Conflicts of Interest

The authors declare no conflicts of interest.

## Authors' Contributions

All authors have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

## Supplementary Materials

The data used to support the findings of this study are provided in "Annexure." (*Supplementary Materials*)

## References

- [1] Z. Chen, L. Li, M. Siew, and T. Q. Quek, *Edge/Fog Computing Networks*, pp. 1–18, Wiley, Hoboken, NJ, USA, 2019.
- [2] H. Tayyabah, A. K. Wajihah, M. Rizwan, A. D. Fahad, and S. Usman, "5G radio access networks based ON fog computing: an overview," *Journal of Natural and Applied Sciences Pakistan*, vol. 1, no. 1, pp. 46–51, 2019.
- [3] H. Wei, H. Luo, and Y. Sun, "Mobility-aware service caching in mobile edge computing for internet of things," *Sensors*, vol. 20, no. 3, p. 610, 2020.
- [4] K. Hasan and S.-H. Jeong, "Efficient caching for data-driven IoT applications and fast content delivery with low latency in ICN," *Applied Sciences*, vol. 9, no. 22, p. 4730, 2019.
- [5] Y. Lan, X. Wang, C. Wang, D. Wang, and Q. Li, "Collaborative computation offloading and resource allocation in cache-aided hierarchical edge-cloud systems," *Electronics*, vol. 8, no. 12, p. 1430, 2019.
- [6] Y.-J. Ku, D.-Y. Lin, C.-F. Lee et al., "5G radio access network design with the fog paradigm: confluence of communications and computing," *IEEE Communications Magazine*, vol. 55, no. 4, pp. 46–52, 2017.
- [7] J. Liu, B. Bai, J. Zhang, and K. B. Letaief, "Cache placement in fog-RANs: from centralized to distributed algorithms," *IEEE Transactions on Wireless Communications*, vol. 16, no. 11, pp. 7039–7051, 2017.
- [8] H. Atlam, R. Walters, and G. Wills, "Fog computing and the internet of things: a review," *Big Data and Cognitive Computing*, vol. 2, no. 2, p. 10, 2018.
- [9] T. Xiao, T. Cui, S. M. Islam, and Q. Chen, "Joint content placement and storage allocation based on federated learning in F-RANs," *Sensors*, vol. 21, no. 1, p. 215, 2021.
- [10] M. Peng, S. Yan, K. Zhang, and C. Wang, "Fog-computing-based radio access networks: issues and challenges," *IEEE Network*, vol. 30, no. 4, pp. 46–53, 2016.
- [11] D. Liu, B. Chen, C. Yang, and A. F. Molisch, "Caching at the wireless edge: design aspects, challenges, and future directions," *IEEE Communications Magazine*, vol. 54, no. 9, pp. 22–28, 2016.
- [12] Y. Liu, M. Peng, G. Shou, Y. Chen, and S. Chen, "Toward edge intelligence: multiaccess edge computing for 5G and internet of things," *IEEE Internet of Things Journal*, vol. 7, no. 8, pp. 6722–6747, 2020.

- [13] S. Yan, M. Jiao, Y. Zhou, M. Peng, and M. Daneshmand, "Machine-learning approach for user association and content placement in fog radio access networks," *IEEE Internet of Things Journal*, vol. 7, no. 10, pp. 9413–9425, 2020.
- [14] G. Qiao, S. Leng, S. Maharjan, Y. Zhang, and N. Ansari, "Deep reinforcement learning for cooperative content caching in vehicular edge computing and networks," *IEEE Internet of Things Journal*, vol. 7, no. 1, pp. 247–257, 2020.
- [15] Y. Jiang, M. Ma, M. Bennis, F.-C. Zheng, and X. You, "User preference learning-based edge caching for fog radio access network," *IEEE Transactions on Communications*, vol. 67, no. 2, pp. 1268–1283, 2019.
- [16] X. Wang, Y. Han, C. Wang, Q. Zhao, X. Chen, and M. Chen, "In-edge AI: intelligentizing mobile edge computing, caching and communication by federated learning," *IEEE Network*, vol. 33, no. 5, pp. 156–165, 2019.
- [17] E. Li, Z. Zhou, and X. Chen, "Edge intelligence: on-demand deep learning model co-inference with device-edge synergy," in *Proceedings of the 2018 Workshop on Mobile Edge Communications*, pp. 31–36, 2018.
- [18] W. L. Chang, L. M. Pang, and K. M. Tay, "Application of self-organizing map to failure modes and effects analysis methodology," *Neurocomputing*, vol. 249, pp. 314–320, 2017.
- [19] J. Faigl, "An application of self-organizing map for multirobot multigoal path planning with Minmax objective," *Computational Intelligence and Neuroscience*, vol. 2016, Article ID 2720630, 15 pages, 2016.
- [20] M. Bazarghan, "Application of self-organizing map to stellar spectral classifications," *Astrophysics and Space Science*, vol. 337, no. 1, pp. 93–98, 2012.
- [21] U. Rohrmann and R. Schäfer, "Stern-gerlach experiments on Fe@Sn12: magnetic response of a jahn-teller distorted endohedrally doped molecular cage cluster," *The Journal of Physical Chemistry C*, vol. 119, no. 20, pp. 10958–10961, 2015.
- [22] X. Xu and Z. Xiao-Ji, "Phase-dependent effects in stern-gerlach experiments," *Chinese Physics Letters*, vol. 27, no. 1, Article ID 010309, 2010.
- [23] T. Hassan, W. A. Khan, F. Ahmad, M. Rizwan, and R. Rehman, "Edge caching framework in fog based radio access networks through AI in quantum regime," in *Communications in Computer and Information Science Intelligent Technologies and Applications*, pp. 711–722, Springer, Berlin, Germany, 2020.
- [24] C.-X. Wang, M. D. Renzo, S. Stanczak, S. Wang, and E. G. Larsson, "Artificial intelligence enabled wireless networking for 5G and beyond: recent advances and future challenges," *IEEE Wireless Communications*, vol. 27, no. 1, pp. 16–23, 2020.
- [25] Z. Chang, L. Lei, Z. Zhou, S. Mao, and T. Ristaniemi, "Learn to cache: machine learning for network edge caching in the big data era," *IEEE Wireless Communications*, vol. 25, no. 3, pp. 28–35, 2018.
- [26] Y. Dai, D. Xu, S. Maharjan, G. Qiao, and Y. Zhang, "Artificial intelligence empowered edge computing and caching for internet of vehicles," *IEEE Wireless Communications*, vol. 26, no. 3, pp. 12–18, 2019.
- [27] M. Usama, J. Qadir, A. Raza et al., "Unsupervised machine learning for networking: techniques, applications and research challenges," *IEEE Access*, vol. 7, pp. 65579–65615, 2019.
- [28] M. McClellan, C. Cervelló-Pastor, and S. Sallent, "Deep learning at the mobile edge: opportunities for 5G networks," *Applied Sciences*, vol. 10, no. 14, p. 4735, 2020.
- [29] Y. Wang and V. Friderikos, "A survey of deep learning for data caching in edge network," *Informatics*, vol. 7, no. 4, p. 43, 2020.
- [30] G. Arun and V. Mishra, "A review on quantum computing and communication," in *Proceedings of the 2014 2nd International Conference on Emerging Technology Trends in Electronics, Communication and Networking*, pp. 1–5, Surat, India, December 2014.
- [31] D. D. Thaker, T. S. Metodi, A. W. Cross, I. L. Chuang, and F. T. Chong, "Quantum memory hierarchies: efficient designs to match available parallelism in quantum computing," in *Proceedings of the 33rd International Symposium on Computer Architecture (ISCA'06)*, pp. 378–390, Boston, MA, USA, June 2006.
- [32] M. A. Sillanpää, J. I. Park, and R. W. Simmonds, "Coherent quantum state storage and transfer between two phase qubits via a resonant cavity," *Nature*, vol. 449, no. 7161, pp. 438–442, 2007.
- [33] R. Vignesh and P. G. Poonacha, "Quantum computer architectures: an idea whose time is not far away," in *Proceedings of the 2015 International Conference on Computers, Communications, and Systems (ICCCS)*, pp. 6–13, Kanyakumari, India, November 2015.
- [34] H. Mäkelä and A. Messina, "N-qubit states as points on the Bloch sphere," *Physica Scripta*, vol. T140, Article ID 014054, 2010.
- [35] O. Giraud, D. Braun, D. Baguette, T. Bastin, and J. Martin, "Tensor representation of spin states," *Physical Review Letters*, vol. 114, no. 8, Article ID 080401, 2015.
- [36] V. Hahanov, S. Chumachenko, E. Litvinova, and H. Khakhanova, "Architectures of quantum memory-driven computing," in *Proceedings of the 2018 IEEE East-West Design & Test Symposium (EWDTS)*, pp. 1–7, Kazan, Russia, September 2018.
- [37] H. Häffner, C. F. Roos, and R. Blatt, "Quantum computing with trapped ions," *Physics Reports*, vol. 469, no. 4, pp. 155–203, 2008.
- [38] D. Konar, S. Bhattacharyya, B. K. Panigrahi, and M. K. Ghose, "An efficient pure color image denoising using quantum parallel bidirectional self-organizing neural network architecture," in *Quantum Inspired Computational Intelligence*, pp. 149–205, Elsevier, Amsterdam, Netherlands, 2017.
- [39] S. Chen, Z. Yao, X. Jiang, J. Yang, and L. Hanzo, "Multi-agent deep reinforcement learning-based cooperative edge caching for ultra-dense next-generation networks," *IEEE Transactions on Communications*, vol. 69, no. 4, pp. 2441–2456, 2021.
- [40] M. Yan, W. Li, C. A. Chan, and S. Bian, "PECS: towards personalized edge caching for future service-centric networks," *China Communications*, vol. 16, no. 8, pp. 93–106, 2019.
- [41] S. A. Alanazi, M. Alruwaili, F. Ahmad, A. Alaerjan, and N. Alshammari, "Estimation of organizational competitiveness by a hybrid of one-dimensional convolutional neural networks and self-organizing maps using physiological signals for emotional analysis of employees," *Sensors*, vol. 21, no. 11, p. 3760, 2021.
- [42] S. Shahzadi, F. Ahmad, A. Basharat et al., "Machine learning empowered security management and quality of service provision in SDN-NFV environment," *Computers, Materials & Continua*, vol. 66, no. 3, pp. 2723–2749, 2021.

## Retraction

# Retracted: Sliding Mode Control of Flexible Articulated Manipulator Based on Robust Observer

### Computational Intelligence and Neuroscience

Received 15 August 2023; Accepted 15 August 2023; Published 16 August 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] Y. Gao and H. Lu, "Sliding Mode Control of Flexible Articulated Manipulator Based on Robust Observer," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 2440770, 10 pages, 2022.



## Research Article

# Sliding Mode Control of Flexible Articulated Manipulator Based on Robust Observer

Yanghua Gao  and Hailiang Lu 

Information Center, China Tobacco Zhejiang Industrial Co., Ltd, Hangzhou 310008, China

Correspondence should be addressed to Yanghua Gao; yhgao@zju.edu.cn

Received 19 November 2021; Revised 14 December 2021; Accepted 16 December 2021; Published 4 January 2022

Academic Editor: Daqing Gong

Copyright © 2022 Yanghua Gao and Hailiang Lu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, a robust observer-based sliding mode control algorithm is proposed to address the modelling and measurement inaccuracies, load variations, and external disturbances of flexible articulated manipulators. Firstly, a sliding mode observer was designed with exponential convergence to observe system state accurately and to overcome the measuring difficulty of the state variables, unmeasurable quantities, and external disturbances. Next, a robust sliding mode controller was developed based on the observer, such that the output error of the system converges to zero in finite time. In this way, the whole system achieves asymptotic stability. Finally, the convergence conditions of the observer were theoretically analyzed to verify the convergence of the proposed algorithm, and simulation was carried out to confirm the effectiveness of the proposed method.

## 1. Introduction

Flexible manipulators are increasingly applied in industrial and aerospace fields, such as welding robots, industrial production lines, mechanical arms of aircraft, and so on, owing to their energy efficiency, high speed, and low contact impact. More and more attention has been paid to the research of flexible manipulators, along with the development of aerospace technology, robotics, marine engineering, and industrial engineering. Flexible manipulators are now extensively used to comfort humans in different areas of work, which involves risky and tedious works such as painting, cutting, dispensing, material handling, machine tending, machining, and assembly. However, each flexible manipulator is an extremely complex, dynamic system with highly nonlinear, strongly coupled, and time-varying features. The system behaviors are complex and dynamic due to load variations, uncertain external perturbations, and inherent vibrations [1–3]. Hence, flexible manipulators can hardly be modelled or measured accurately, calling for a well-designed controller [4–8]. Against this backdrop, it is theoretically and practically significant to explore the response speed and control accuracy of trajectory tracking for the double-linked flexible-joint manipulator [9].

To address the above problems, Lee et al. [10] designed an adaptive proportional-derivative (PD) controller to improve the trajectory tracking accuracy of the flexible-joint manipulator but did not consider the stability of the manipulator system. Lee and Lee [11] proposed a hybrid control strategy to optimize the design of the controller and generate hybrid trajectories. The strategy enhances the robustness of the flexible-joint manipulator system, yet it failed to take into account the trajectory tracking accuracy of the manipulator. Dong et al. [12] presented a fuzzy optimal control method for the design of a robust adaptive controller and demonstrated that the method ensures accurate and robust trajectory tracking of the flexible-joint manipulator. However, the manipulator's response speed of trajectory tracking was not taken into consideration. Abd Latip et al. [13] automatically adjusted the control gain online with an adaptive proportional-integral-derivative (PID) controller, which supports the control of the single-link flexible manipulator even after the actuator failure.

Ahanda et al. [14] addressed the robust adaptive control of a robotic manipulator under uncertain dynamics and joint space constraints and adopted command filters to overcome the time derivatives of virtual control, eliminating



the need for differentiating the desired trajectory. In addition, a barrier Lyapunov function was introduced to handle joint space constraints, and a robust adaptive support vector regression architecture was employed to suppress filtering errors, approximation errors, and dynamic uncertainties. Based on unknown input observer (UIO), Wang et al. [15] put forward a novel funnel nonsingular terminal sliding mode control (FNTSMC) method for servomechanisms with unknown dynamics, e.g., nonlinear friction, uncertainties, and external disturbances. He et al. [16] created a full-state feedback neural network (NN) control to mitigate the uncertainties and enhance the robustness of the dynamic system of a flexible-joint manipulator. Through a Lyapunov stability analysis, it was demonstrated that the controller can ensure the stability of the flexible-joint manipulator system and guarantee the boundedness of system state variables, by choosing appropriate control gains. Rahmani and Belkheiri [17] came up with a novel approach for adaptive control of flexible multilink robots in the joint space, proved that the approach is valid for a class of highly uncertain systems with arbitrary but bounded dimensions, and realized trajectory tracking by developing a stable inversion for robot dynamics using only joint angle measurements. Guo et al. [18] investigated the repetitive motion planning (RMP) of robotic manipulators under the high precision of joint angle repeatability and end-effector motion and applied a special difference rule to discretize the existing RMP scheme with  $P$ -based formulation, yielding a novel pseudoinverse-based ( $P$ -based) RMP scheme for robotic manipulators.

Through the above analysis, this paper proposes a sliding mode control strategy based on the robust observer. Firstly, a sliding mode robust observer was designed in light of the unmeasurable state, the modelling uncertainty, and the external disturbance moment of the flexible-joint manipulator. Next, a sliding mode controller was designed to track the positions of the first and second joints of the manipulator, aiming to realize the finite-time control of the system. Meanwhile, the convergence of the observer and controller was analyzed to present the convergence conditions. Finally, the effectiveness of the proposed method was verified through simulation.

## 2. Problem Description

The dynamics of the flexible-joint manipulator can be expressed as

$$\begin{cases} I\ddot{\theta} + K(\theta - \theta_m) + Mgl \sin q = 0, \\ J\ddot{\theta}_m - K(\theta - \theta_m) = u, \end{cases} \quad (1)$$

where  $\theta$  and  $\theta_m$  are the angular positions of the link and rotor, respectively;  $I$  and  $J$  represent the rotational inertia of the link and rotor, respectively;  $K$  is the joint stiffness coefficient;  $M$ ,  $g$ , and  $l$  are the link mass, gravitational acceleration at the link's center of gravity, and joint length, respectively; and  $u$  is the motor torque input.

Let  $x_1 = \theta$ ,  $x_2 = \dot{\theta}$ ,  $x_3 = \theta_m$ , and  $x_4 = \dot{\theta}_m$  be state variables. Considering modelling uncertainty and external disturbance moments, the underdriven form of equation (1) can be obtained as

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = a_1 x_3 + f_1(x_1) + \Delta_1(t), \\ \dot{x}_3 = x_4, \\ \dot{x}_4 = a_2 u + f_2(x_1, x_3) + \Delta_2(t), \end{cases} \quad (2)$$

where  $a_1 = (K/I)$ ;  $f_1(x_1) = -(Mgl/I) \cdot \sin x_1 - (K/I) \cdot x_1$ ;  $a_2 = (1/J)$ ;  $f_2(x_1, x_3) = (K/J) \cdot (x_1 - x_3)$ ;  $\Delta_1(t)$  and  $\Delta_2(t)$  are the uncertainty part and the external disturbance moment, respectively; and  $|\Delta_1(t)| \leq \rho_1$ ,  $|\Delta_2(t)| \leq \rho_2$ .

The following lemma was introduced to facilitate the observer and controller stability analysis.

**Lemma 1** (see [1]). *For  $V: [0, \infty) \in R$ , the solution of  $\dot{V} \leq -\alpha V + f$  with  $\forall t \geq t_0 \geq 0$  can be expressed as an inequality:*

$$V(t) \leq e^{-\alpha(t-t_0)}V(t_0) + \int_{t_0}^t e^{-\alpha(t-\tau)}f(\tau)d\tau, \quad (3)$$

where  $\alpha$  is an arbitrary constant.

## 3. The Observer and Controller Design

**3.1. The Observer Design.** The observer of  $x_2$  and  $x_4$  was designed as follows.

To realize  $x_2$  and  $x_4$  observations, the following reconfiguration system was developed:

$$\begin{cases} \dot{\lambda}_1 = \lambda_2 + l_1(x_1 - \lambda_1) + D_1(x_1 - \lambda_1), \\ \dot{\lambda}_2 = a_1 x_3 + f_1(x_1) + \overline{D}_2(x_1 - \lambda_1), \\ \dot{\lambda}_3 = \lambda_1 + l_2(x_3 - \lambda_3) + D_3(x_3 - \lambda_3), \\ \dot{\lambda}_4 = a_2 u + f_2(x_1, x_3) + \overline{D}_4(x_3 - \lambda_3), \end{cases} \quad (4)$$

where  $l_1, l_2, D_1, \overline{D}_2, D_3$ , and  $\overline{D}_4$  are the positive real numbers to be designed and  $\lambda_1, \lambda_2, \lambda_3$ , and  $\lambda_4$  are meaningless intermediate state variables.

Then, the observer was designed as

$$\begin{cases} \hat{x}_1 = \lambda_1, \\ \hat{x}_2 = \lambda_2 + l_1(x_1 - \lambda_1), \\ \hat{x}_3 = \lambda_3, \\ \hat{x}_4 = \lambda_4 + l_2(x_3 - \lambda_3), \end{cases} \quad (5)$$

where  $\hat{x}_i$  is the state estimation. The estimation error can be defined as

$$\tilde{x}_i = x_i - \hat{x}_i. \quad (6)$$

From equations (4)–(6), we have

$$\begin{cases} \dot{\hat{x}}_1 = \lambda_2 + l_1(x_1 - \lambda_1) + D_1(x_1 - \lambda_1) = \hat{x}_2 + D_1\tilde{x}_1, \\ \dot{\hat{x}}_2 = a_1x_3 + f_1(x_1) + \bar{D}_2(x_1 - \lambda_1) + l_1(x_2 - \hat{x}_2 - D_1\tilde{x}_1) \\ = a_1x_3 + f_1(x_1) + l_1\tilde{x}_2 + (\bar{D}_2 - l_1D_1)\tilde{x}_1, \\ \dot{\hat{x}}_3 = \lambda_4 + l_2(x_3 - \lambda_3) + D_3(x_3 - \lambda_3) = \hat{x}_4 + D_3\tilde{x}_3, \\ \dot{\hat{x}}_4 = a_2u + f_2(x_1, x_3) + \bar{D}_4(x_3 - \lambda_3) + l_2(x_4 - \hat{x}_4 - D_3\tilde{x}_3) \\ = a_2u + f_2(x_1, x_3) + l_2\tilde{x}_4 + (\bar{D}_4 - l_2D_3)\tilde{x}_3. \end{cases} \quad (7)$$

Note that  $D_2 = \bar{D}_2 - l_1D_1$  and  $D_4 = \bar{D}_4 - l_2D_3$ . Then,

$$\begin{cases} \dot{\hat{x}}_1 = \hat{x}_2 + D_1\tilde{x}_1, \\ \dot{\hat{x}}_2 = a_1x_3 + f_1(x_1) + l_1\tilde{x}_2 + D_2\tilde{x}_1, \\ \dot{\hat{x}}_3 = \hat{x}_4 + D_3\tilde{x}_3, \\ \dot{\hat{x}}_4 = a_2u + f_2(x_1, x_3) + l_2\tilde{x}_4 + D_4\tilde{x}_3. \end{cases} \quad (8)$$

$$\begin{aligned} \dot{V} &= \tilde{x}_1(x_2 - \hat{x}_2 - D_1\tilde{x}_1) + \tilde{x}_2(\Delta_1 - l_1\tilde{x}_2 - D_2\tilde{x}_1) + \tilde{x}_3(x_4 - \hat{x}_4 - D_3\tilde{x}_3) + \tilde{x}_4(\Delta_2 - l_2\tilde{x}_4 - D_1\tilde{x}_3) \\ &= (1 - D_2)\tilde{x}_1\tilde{x}_2 + (1 - D_4)\tilde{x}_3\tilde{x}_4 - D_1\tilde{x}_1^2 - l_1\tilde{x}_2^2 - D_3\tilde{x}_3^2 - l_2\tilde{x}_4^2 + \Delta_1\tilde{x}_2 + \Delta_2\tilde{x}_4. \end{aligned} \quad (10)$$

Taking  $D_2 = D_1 = 1$  and the inequality  $\rho_i^2/2 + \tilde{x}_j^2/2 \geq \rho_i |\tilde{x}_j| \geq \Delta_i \tilde{x}_i$ , we have

$$\dot{V} \leq -D_1\tilde{x}_1^2 - l_1\tilde{x}_2^2 - D_3\tilde{x}_3^2 - l_2\tilde{x}_4^2 + \frac{\rho_1^2}{2} + \frac{\tilde{x}_2^2}{2} + \frac{\rho_2^2}{2} + \frac{\tilde{x}_4^2}{2}. \quad (11)$$

Inequality (11) can be rectified as

$$\dot{V} \leq -\left(D_1\tilde{x}_1^2 + D_3\tilde{x}_3^2 + \left(l_1 - \frac{1}{2}\right)\tilde{x}_2^2 + \left(l_2 - \frac{1}{2}\right)\tilde{x}_4^2\right) + \frac{\rho_1^2}{2} + \frac{\rho_2^2}{2}. \quad (12)$$

Taking  $l_1 \geq (1/2) + r$ ,  $l_2 \geq (1/2) + r$ ,  $D_1 \geq r$ , and  $D_3 \geq r$  with  $r$  being the positive real number to be designed,

$$V(t) \leq e^{-2r(t-t_0)}V(t_0) + Qe^{-2rt} \int_{t_0}^t e^{2r\tau} d\tau = e^{-2r(t-t_0)}V(t_0) + \frac{Qe^{-2rt}}{2r} (e^{2rt} - e^{2rt_0}) = e^{-2r(t-t_0)}V(t_0) + \frac{Q}{2r} (1 - e^{-2r(t-t_0)}). \quad (15)$$

That is,

$$V(t) \leq \frac{Q}{2r} + \left(V(t_0) - \frac{Q}{2r}\right)e^{-2r(t-t_0)}. \quad (16)$$

Then, all the signals of the system are semiglobally bounded and satisfy

$$\lim_{t \rightarrow \infty} V(t) \leq \frac{Q}{2r}. \quad (17) \quad \square$$

*Remark 1.* From equation (17), it can be inferred that the observation accuracy of the state depends on the upper bound  $\Delta_1(t)$  and the initial error  $\Delta_2(t)$  of the observer.

The following theorem was introduced to facilitate the proof of observer convergence.

**Theorem 1.** For system (2) and observer (5), if the initial conditions satisfy  $V(0) \leq p$ , where  $p$  is any positive real number, there exists a condition that all the signals  $l_1, l_2, D_i (i = 1, \dots, 4)$  of the system are semiglobally consistent and bounded, and the observation error converges to an arbitrarily small residual set.

*Proof.* According to equations (5) and (6), the Lyapunov function is taken as

$$V = \frac{1}{2} \sum_{i=1}^4 \tilde{x}_i^2. \quad (9)$$

The following can be derived from equation (9):

$$D_1\tilde{x}_1^2 + D_3\tilde{x}_3^2 + \left(l_1 - \frac{1}{2}\right)\tilde{x}_2^2 + \left(l_2 - \frac{1}{2}\right)\tilde{x}_4^2 \geq r(\tilde{x}_1^2 + \tilde{x}_3^2 + \tilde{x}_2^2 + \tilde{x}_4^2). \quad (13)$$

Thus,

$$\dot{V} \leq -r(\tilde{x}_1^2 + \tilde{x}_3^2 + \tilde{x}_2^2 + \tilde{x}_4^2) + \frac{\rho_1^2}{2} + \frac{\rho_2^2}{2} \leq -2rV + Q, \quad (14)$$

where  $Q = \rho_1^2/2 + \rho_2^2/2$ .

According to Lemma 1, the solution to inequality (14) is

When parameter  $r$  is infinitely large, the observation error will be arbitrarily small.

*Remark 2.* Without considering the modelling uncertainty  $\Delta_1(t) = 0$  and the external disturbance moment  $\Delta_2(t) = 0$ ,  $(\rho_1^2/2) + (\rho_2^2/2) = 0$ , that is, if  $\dot{V} \leq -2rV$ , then  $V(t) \leq e^{-2r(t-t_0)}V(t_0)$ . At this point, the observer converges exponentially.

**3.2. Design and Analysis of Observer-Based Sliding Mode Controller.** Observer-based sliding mode control is a new sliding mode control method in recent years. It solves the unknown disturbance problem directly from the sliding

mode design side by purposefully designing the switching function and realizes the global nonsingular control of the system. At the same time, it inherits the finite-time convergence characteristics of sliding mode. Compared with the traditional sliding mode control, it can make the control system converge to the desired trajectory in finite time and has high steady-state accuracy. It is especially suitable for high-speed and high-precision control.

Let  $x_1$  and  $x_2$  be the controlled targets of  $x_d$  and  $\dot{x}_d$ , respectively. The design error can be expressed as

$$\begin{cases} e_1 = x_1 - x_d, \\ e_2 = \dot{e}_1 = x_2 - \dot{x}_d, \\ e_3 = \ddot{e}_1 = \dot{x}_2 - \ddot{x}_d = a_1 x_3 + f_1(x_1) - \ddot{x}_d, \\ e_4 = \ddot{\ddot{e}}_1 = a_1 \dot{x}_3 + \dot{f}_1(x_1) - \dot{\ddot{x}}_d = a_1 x_4 + \dot{f}_1(x_1) - \dot{\ddot{x}}_d. \end{cases} \quad (18)$$

Then, the error of the observer can be expressed as

$$\hat{s} = c_1 \tilde{e}_1 + c_2 \tilde{e}_2 + c_3 \tilde{e}_3 + \tilde{e}_4 = c_1 \tilde{x}_1 + c_2 \tilde{x}_2 + c_3 (a_1 \tilde{x}_3 + f_1(x_1) - \hat{f}_1(x_1)) + a_1 \tilde{x}_4 + \dot{f}_1(x_1) - \hat{f}_1(x_1). \quad (22)$$

Then, the control law can be designed as

$$u = -\frac{1}{a_1 a_2} \left[ c_1 (\tilde{x}_2 - \dot{x}_d) + c_2 (a_1 \tilde{x}_3 + \hat{f}_1(x_1) - \ddot{x}_d) + c_3 (a_1 \tilde{x}_4 + \dot{f}_1(x_1) - \dot{\ddot{x}}_d) + a_1 \hat{f}_1(x_1, x_3) + \hat{f}_1(x_1) - \dot{\ddot{x}}_d + \eta \hat{s} \right], \quad (23)$$

where  $\eta > 0$ .

$$\begin{cases} \hat{e}_1 = \hat{x}_1 - x_d, \\ \hat{e}_2 = \hat{x}_2 - \dot{x}_d, \\ \hat{e}_3 = a_1 \hat{x}_3 + \hat{f}_1(x_1) - \ddot{x}_d, \\ \hat{e}_4 = a_1 \hat{x}_4 + \hat{f}_1(x_1) - \dot{\ddot{x}}_d. \end{cases} \quad (19)$$

From equations (18) and (19), we have

$$\begin{cases} \tilde{e}_1 = e_1 - \hat{e}_1 = \tilde{x}_1, \\ \tilde{e}_2 = e_2 - \hat{e}_2 = \tilde{x}_2, \\ \tilde{e}_3 = e_3 - \hat{e}_3 = a_1 \tilde{x}_3 + f_1(x_1) - \hat{f}_1(x_1), \\ \tilde{e}_4 = e_4 - \hat{e}_4 = a_1 \tilde{x}_4 + \dot{f}_1(x_1) - \hat{f}_1(x_1). \end{cases} \quad (20)$$

The design sliding mode function can be expressed as

$$s = c_1 e_1 + c_2 e_2 + c_3 e_3 + e_4, \quad (21)$$

where  $c_i > 0$  is designed by  $i = 1, 2, 3$ .

Then,

Then, we have

$$\begin{aligned} \dot{s} &= c_1 \dot{e}_1 + c_2 \dot{e}_2 + c_3 \dot{e}_3 + \dot{e}_4 \\ &= c_1 (x_2 - \dot{x}_d) + c_2 (a_1 x_3 + f_1(x_1) - \dot{x}_d) + c_3 (a_1 x_4 + \dot{f}_1(x_1) - \dot{\ddot{x}}_d) + a_1 (a_2 u + f_1(x_1, x_3)) + \ddot{f}_1(x_1) - \dot{\ddot{x}}_d \\ &= c_1 (x_2 - \dot{x}_d) + c_2 (a_1 x_3 + f_1(x_1) - \dot{x}_d) + c_3 (a_1 x_4 + \dot{f}_1(x_1) - \dot{\ddot{x}}_d) \\ &\quad - \left[ c_1 (\tilde{x}_2 - \dot{x}_d) + c_2 (a_1 \tilde{x}_3 + \hat{f}_1(x_1) - \dot{x}_d) + c_3 (a_1 \tilde{x}_4 + \hat{f}_1(x_1) - \dot{\ddot{x}}_d) + a_1 \hat{f}_1(x_1, x_3) + \hat{f}_1(x_1) - \dot{\ddot{x}}_d + \eta \hat{s} \right] \\ &\quad + a_1 f_1(x_1, x_3) + \ddot{f}_1(x_1) - \dot{\ddot{x}}_d \\ &= c_1 \tilde{x}_2 + c_2 a_1 \tilde{x}_3 + c_3 a_1 \tilde{x}_4 + c_2 (f_1(x_1) - \hat{f}_1(x_1)) \\ &\quad + c_3 (\dot{f}_1(x_1) - \hat{f}_1(x_1)) + a_1 (f_1(x_1, x_3) - \hat{f}_1(x_1, x_3)) + (\ddot{f}_1(x_1) - \hat{f}_1(x_1)) - \eta (s - \hat{s}). \end{aligned} \quad (24)$$

Take the Lyapunov function as

$$V_c = \frac{1}{2}s^2. \quad (25)$$

Then,

$$\begin{aligned} \dot{V}_c &= s\dot{s} = s \left[ c_1\tilde{x}_2 + c_2a_1\tilde{x}_3 + c_3a_1\tilde{x}_4 + c_2(f_1(x_1) - \hat{f}_1(x_1)) + c_3(\dot{f}_1(x_1) - \hat{\dot{f}}_1(x_1)) \right. \\ &\quad \left. + a_1(f_1(x_1, x_3) - \hat{f}_1(x_1, x_3)) + (\ddot{f}_1(x_1) - \hat{\ddot{f}}_1(x_1)) - \eta(s - \bar{s}) \right] \\ &= -\eta s^2 + \eta s \left( c_1\tilde{x}_1 + c_2\tilde{x}_2 + c_3(a_1\tilde{x}_3 + f_1(x_1) - \hat{f}_1(x_1)) + a_1\tilde{x}_4 + \dot{f}_1(x_1) - \hat{\dot{f}}_1(x_1) \right) \\ &\quad + s \left[ c_1\tilde{x}_2 + c_2a_1\tilde{x}_3 + c_3a_1\tilde{x}_4 + c_2(f_1(x_1) - \hat{f}_1(x_1)) + c_3(\dot{f}_1(x_1) - \hat{\dot{f}}_1(x_1)) \right. \\ &\quad \left. + a_1(f_1(x_1, x_3) - \hat{f}_1(x_1, x_3)) + (\ddot{f}_1(x_1) - \hat{\ddot{f}}_1(x_1)) \right] \\ &= -\eta s^2 + s \left[ \eta c_1\tilde{x}_1 + (\eta c_2 + c_1)\tilde{x}_2 + (\eta c_3a_1 + c_2a_1)\tilde{x}_3 + (\eta a_1 + c_3a_1)\tilde{x}_4 \right. \\ &\quad \left. + (\eta c_3 + c_2)(f_1(x_1) - \hat{f}_1(x_1)) + (\eta + c_3)(\dot{f}_1(x_1) - \hat{\dot{f}}_1(x_1)) \right. \\ &\quad \left. + a_1(f_1(x_1, x_3) - \hat{f}_1(x_1, x_3)) + (\ddot{f}_1(x_1) - \hat{\ddot{f}}_1(x_1)) \right] \\ &= -\eta s^2 + s\chi(\tilde{x}) \leq -\eta s^2 + \frac{1}{2}s^2 + \frac{1}{2}\chi^2(\tilde{x}) \\ &= \left(\frac{1}{2} - \eta\right)s^2 + \frac{1}{2}\chi^2(\tilde{x}) = (1 - 2\eta)V_c + \frac{1}{2}\chi^2(\tilde{x}), \end{aligned} \quad (26)$$

where

$$\begin{aligned} \chi(\tilde{x}) &= \left[ \eta c_1\tilde{x}_1 + (\eta c_2 + c_1)\tilde{x}_2 + (\eta c_3a_1 + c_2a_1)\tilde{x}_3 + (\eta a_1 + c_3a_1)\tilde{x}_4 \right. \\ &\quad \left. + (\eta c_3 + c_2)(f_1(x_1) - \hat{f}_1(x_1)) + (\eta + c_3)(\dot{f}_1(x_1) - \hat{\dot{f}}_1(x_1)) \right. \\ &\quad \left. + a_1(f_1(x_1, x_3) - \hat{f}_1(x_1, x_3)) + \ddot{f}_1(x_1) - \hat{\ddot{f}}_1(x_1) \right]. \end{aligned} \quad (27)$$

Since observer (5) converges exponentially, i.e., at time  $t \rightarrow \infty$ ,  $\tilde{x}_1$  converges exponentially to  $\tilde{x}_2$ , and  $\tilde{x}_3$  to  $\tilde{x}_4$ . According to the Taylor series expansion of  $f_1(x_1) = -Mg/I \cdot \sin x_1 - K/I \cdot x_1$  and  $f_2(x_1, x_3) = K/I \cdot (x_1 - x_3)$ ,  $f_1(x_4) \rightarrow \hat{f}_1(x_1)$  converges exponentially to  $\dot{f}_1(x_1) \rightarrow \hat{\dot{f}}_1(x_1)$ ,  $f_1(x_1, x_3) \rightarrow \hat{f}_1(x_1, x_3)$ . Thus,  $\ddot{f}_1(x_1) \rightarrow \hat{\ddot{f}}_1(x_1)$  also converges exponentially to 0.

Considering the observer and the controller, the Lyapunov function of the closed loop is taken as

$$V = V_c + V_o. \quad (28)$$

According to equation (28), we have

$$\begin{aligned} \dot{V} &= \dot{V}_o + \dot{V}_c \leq -2rV_o - (2\eta - 1)V_c \\ &\quad + \frac{1}{2}\chi^2(\tilde{x}) \leq -\eta_1V + \chi(\cdot)e^{-\sigma_0(t-t_0)}, \end{aligned} \quad (29)$$

where  $\eta_1 = \{2r, (2\eta - 1)\}_{\max}$ ;  $\chi(\cdot)$  is the class  $K$  function of  $\|\tilde{x}(t_0)\|$ ; and  $\sigma_0 > 0$ .

According to Lemma 1, the solution to  $\dot{V} \leq -\eta_1V + \chi(\cdot)e^{-\sigma_0(t-t_0)}$  can be expressed as an inequality:

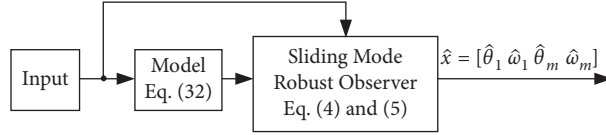
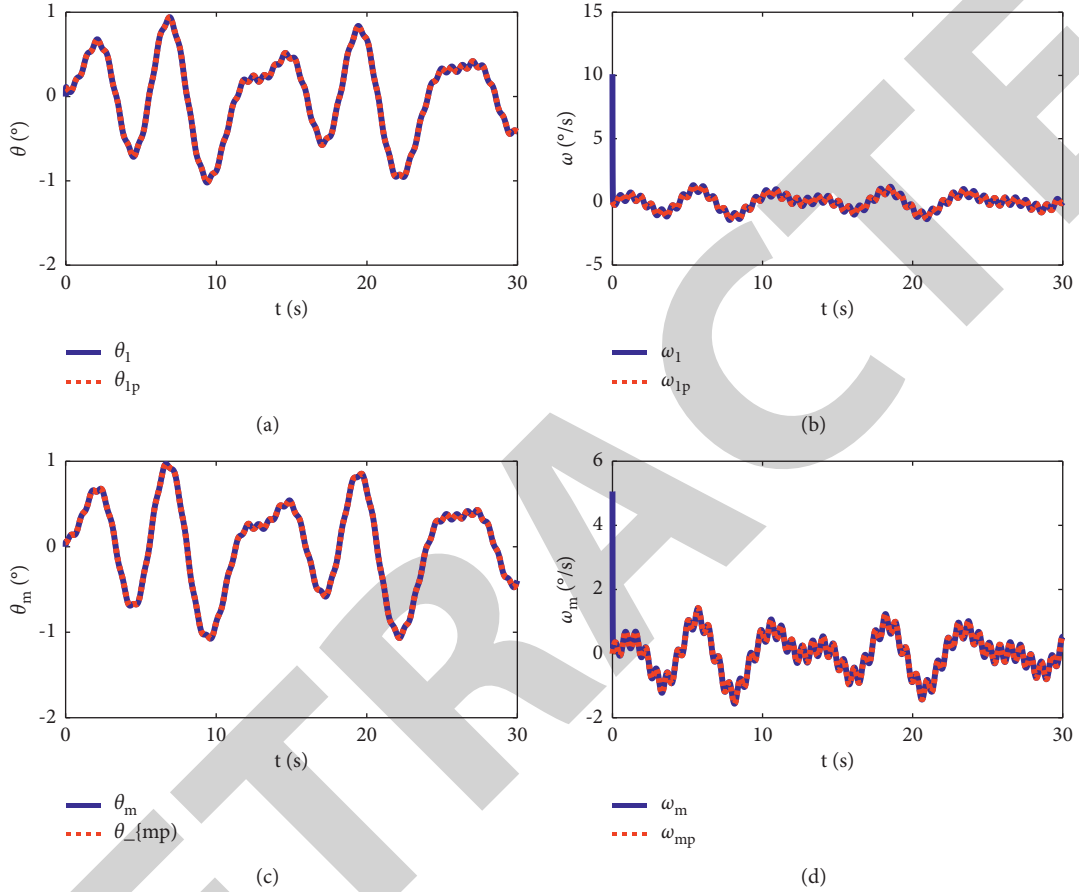


FIGURE 1: Simulation structure of the observer.

FIGURE 2: State estimations. (a)  $\theta$  state estimation. (b)  $\omega$  state estimation. (c)  $\theta_m$  state estimation. (d)  $\omega_m$  state estimation.

$$\begin{aligned}
 V(t) &\leq e^{-\eta_1(t-t_0)}V(t_0) + \chi(\Delta) \int_{t_0}^t e^{-\eta_1(t-\tau)} e^{-\sigma_0(\tau-t_0)} d\tau \\
 &= e^{-\eta_1(t-t_0)}V(t_0) + \chi(\cdot) e^{-\eta_1 t + \sigma_0 t_0} \int_{t_0}^t e^{\eta_1 \tau} e^{-\sigma_0 \tau} d\tau \\
 &= e^{-\eta_1(t-t_0)}V(t_0) + \frac{\chi(\Delta)}{\eta_1 - \sigma_0} e^{-\eta_1 t + \sigma_0 t_0} \left. e^{(\eta_1 - \sigma_0)\tau} \right|_{t_0}^t \\
 &= e^{-\eta_1(t-t_0)}V(t_0) + \frac{\chi(\Delta)}{\eta_1 - \sigma_0} e^{-\eta_1 t + \sigma_0 t_0} \left( e^{(\eta_1 - \sigma_0)t} - e^{(\eta_1 - \sigma_0)t_0} \right) \\
 &= e^{-\eta_1(t-t_0)}V(t_0) + \frac{\chi(\cdot)}{\eta_1 - \sigma_0} \left( e^{-\sigma_0(t-t_0)} - e^{-\eta_1(t-t_0)} \right).
 \end{aligned} \tag{30}$$

That is,  $\lim_{t \rightarrow \infty} V(t) \leq 0$ .

Since  $V(t) \geq 0$ , when  $t \rightarrow \infty$ ,  $V(t) = 0$ , and  $V(t)$  converges exponentially. The convergence accuracy depends on  $\eta_1$ , i.e.,  $r$  and  $\eta$ .

*Remark 3.* When the controller reaches the sliding mode surface, that is,  $s = 0$ , we have  $e_4 = -c_1 e_1 - c_2 e_2 - c_3 e_3$ . If

$$E_1 = [e_1 \ e_2 \ e_3]^T \quad \text{and} \quad A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -c_1 & -c_2 & -c_3 \end{bmatrix}, \quad \text{then}$$

$\dot{E}_1 = AE_1$ . Through the design of  $c_1$ ,  $c_2$ , and  $c_3$ ,  $A$  is Hurwitz zeta function. Thus, at time  $t \rightarrow \infty$ ,  $E_1 = [e_1 \ e_2 \ e_3]^T \rightarrow 0$ . To make  $A$  as a Hurwitz zeta function, the real root part of the following equation must be negative:

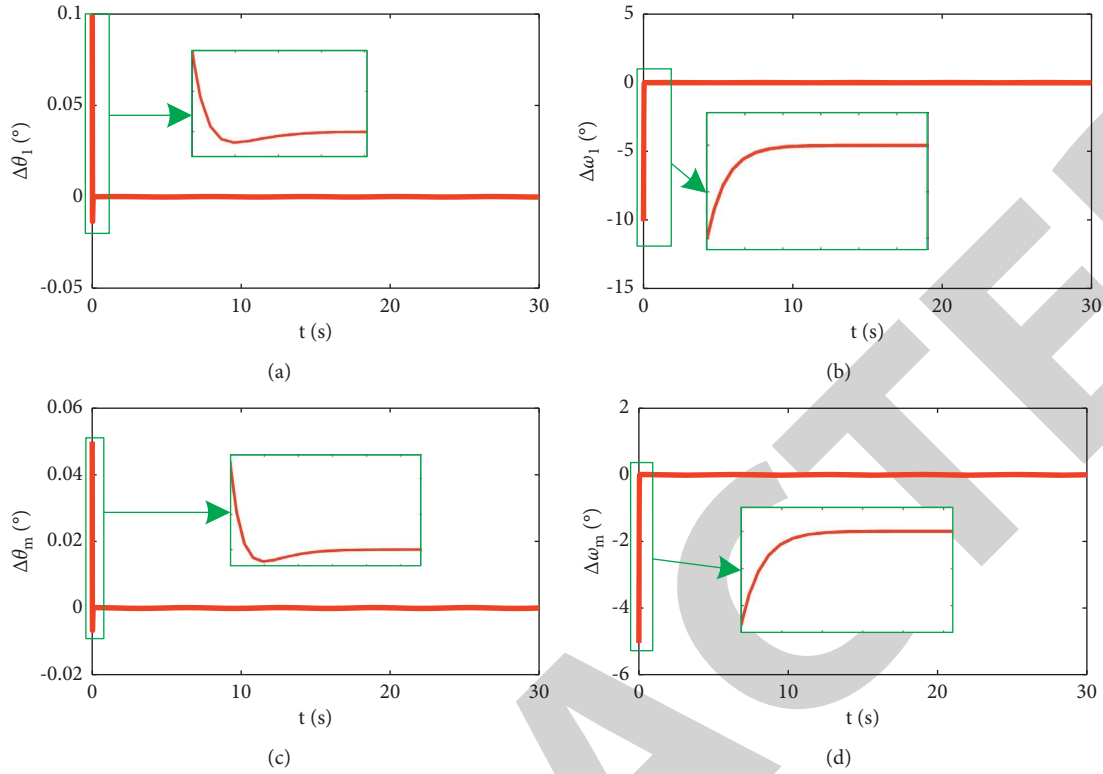
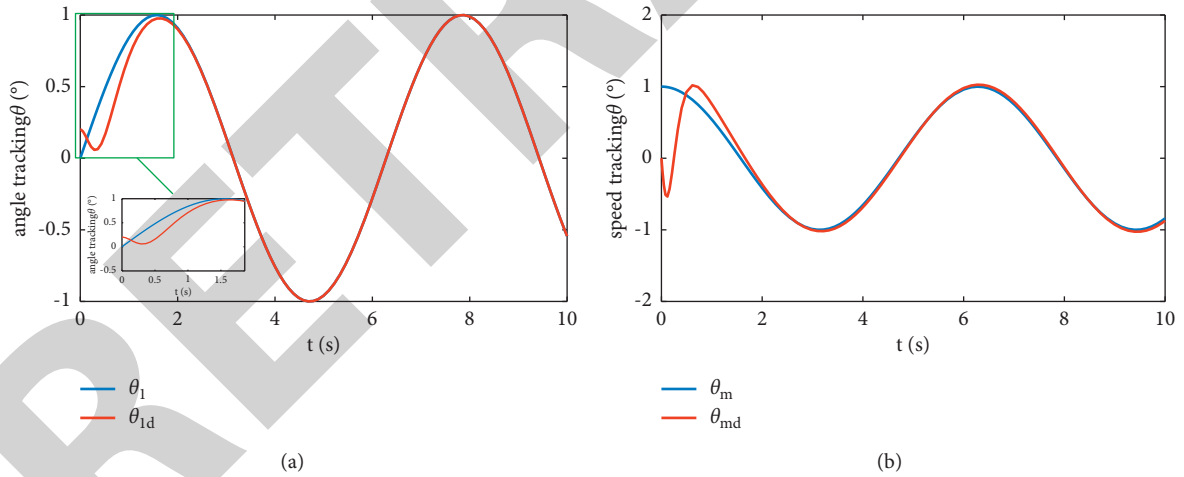

 FIGURE 3: State estimation errors. (a)  $\theta$  estimation error. (b)  $\omega$  estimation error. (c)  $\theta_m$  estimation error. (d)  $\omega_m$  estimation error.


FIGURE 4: Angle and angular velocity tracking with the proposed method. (a) Angle tracking. (b) Angular velocity tracking.

$$|A - \lambda I| = \begin{vmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 1 \\ -c_1 - c_2 & -c_3 - \lambda & \end{vmatrix} = \lambda^2(-c_3 - \lambda) - c_1 - c_2\lambda. \quad (31)$$

That is,  $-\lambda^3 - c_3\lambda^2 - c_2\lambda - c_1 = 0$ . Taking the eigenvalue of  $-10$ ,  $(\lambda + 10)^3 = 0$ , from  $\lambda^3 + 9\lambda^2 + 27\lambda + 27 = 0$ , we can obtain that  $\lambda^3 + c_3\lambda^2 + c_2\lambda + c_1 = 0$ ,  $c_1 = 1000$ ,  $c_2 = 300$ , and  $c_3 = 30$ . Hence, the convergence condition can be satisfied.



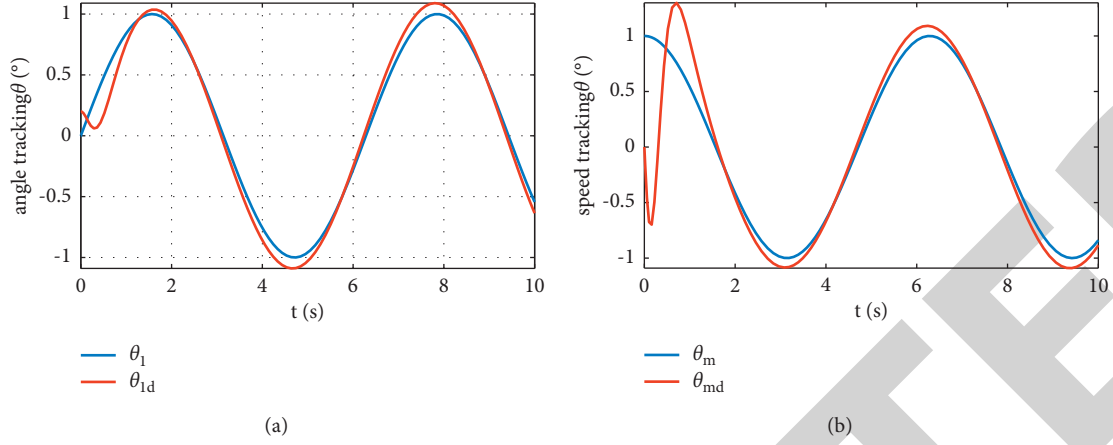


FIGURE 5: Angle and angular velocity tracking with PID control. (a) Angle tracking. (b) Angular velocity tracking.

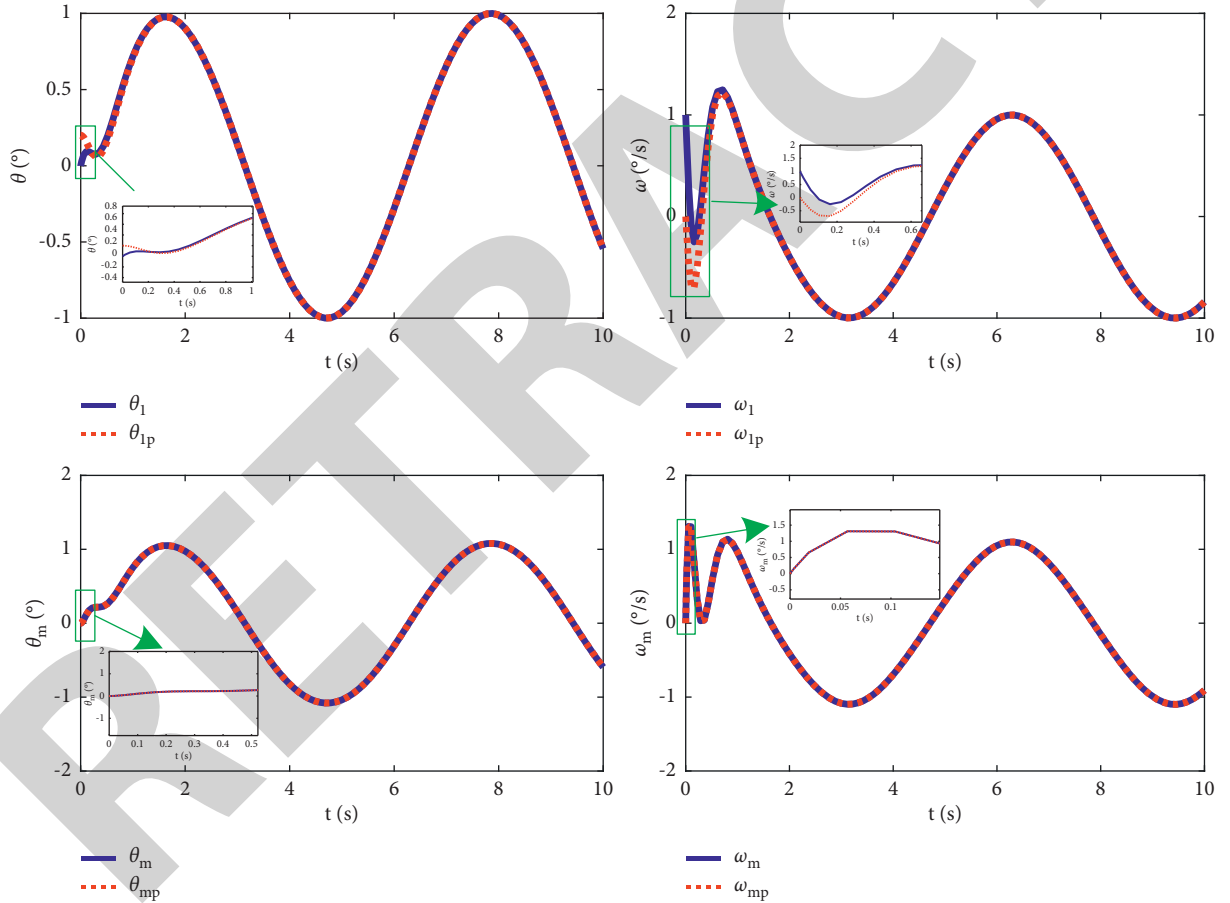


FIGURE 6: Observation of each state of the manipulator.

## 4. Results and Discussion

**4.1. The Simulation of the Observer.** To verify the feasibility of the robust observer, a system was developed to run in an open loop and modelled considering the following external disturbance moments and modelling uncertainties:

$$\begin{cases} \dot{\theta}_1 = \omega_1, \\ \dot{\omega}_1 = a_1\theta_m + f_1(\theta_1) + \Delta_1(t), \\ \dot{\theta}_m = \omega_m, \\ \dot{\omega}_m = a_2u + f_2(\theta_1, \theta_m) + \Delta_2(t), \end{cases} \quad (32)$$

where  $\theta_1$ ,  $\omega_1$ ,  $\theta_m$ , and  $\omega_m$  are the position of rod 1, the angular velocity of rod 1, the position of rod  $m$ , and the angular velocity of rod  $m$ , respectively. The parameters were configured as follows:  $x = [\theta_1 \omega_1, \dots, \theta_m \omega_m]^T$ ,  $\Delta_1(t) = \sin t$ ,  $\Delta_2(t) = \cos t$ ,  $J = 1 \text{ kg} \cdot \text{m}^2$ ,  $Mgl = 5 \text{ Nm}$ , and  $K = 40 \text{ Nm/rad}$ . Before simulation, the system state was initialized as  $x(0) = [0.1 \ 0 \ 0.05 \ 0]^T$ , and the observer state is initialized as  $\lambda(0) = [0 \ 0 \ 0 \ 0]^T$ . The observer adopts the form of (4) and (5), with  $r = 100$ . Based on  $l_1 \geq 1/2 + r$  (5),  $l_2 \geq 1/2 + r$ ,  $D_1 \geq r$ , and  $D_2 \geq r$ , the following parameter values were selected:  $l_1 = l_2 = 101$ ,  $D_2 = D_4 = 1.0$ , and  $D_1 = D_3 = 101$ .

The simulation structure of the observer is given in Figure 1, and the simulation results are shown in Figures 2 and 3. Specifically, Figure 2 presents the flexible modes of the position states of the two joints and their derivatives (i.e., velocities), and Figure 3 displays the tracking errors of the states. It can be inferred that the proposed observer can completely observe each state of the system (as suggested by Figure 2) and fully track the states of the upper two joints after only 0.1 s (as indicated by the error curve in Figure 3, the errors are 0.001, 0.15, 0.0015, and 0.12 for Figures 3(a)–3(d)). Therefore, our method was proved to be fast and effective. Although there are disturbances in the system, i.e.,  $\Delta_1(t) = \sin t$  and  $\Delta_2(t) = \cos t$ , the observation results show the anti-interference ability and good robustness of the proposed observer.

**4.2. The Simulation of the Control Algorithm.** To verify the effectiveness of the proposed control algorithm, the system with  $\Delta_1(t) = 0.15 \sin t$  and  $\Delta_2(t) = 0.23 \cos t$  was taken as shown in equation (32), where  $x(0) = [0.2 \ 0 \ 0 \ 0]^T$  and disturbance torque is  $\lambda(0) = [0 \ 0 \ 0 \ 0]^T$ . The other parameters were kept the same as in simulation 1. The controller takes equation (23), with  $c_1 = 1000$ ,  $c_2 = 300$ ,  $c_3 = 30$ , and  $\eta = 1.5$ . The desired trajectory of joint 2 is  $\theta_{1d} = \sin$ . The simulation results are shown in Figures 4 and 5. The former presents the angle and angular velocity of the second joint of the manipulator, and the latter exhibits the observed values of each state of the manipulator.

As shown in Figure 4, the system state was stabilized in a limited time, despite the presence of external disturbances and fault signals, indicating that the system converges well under this controller. Because the initial state of the system is  $x(0) = [0.2 \ 0 \ 0 \ 0]^T$  and due to the existence of interference, there is a large error at the initial time. However, with the increase of control time, the system error decreases rapidly. Hence, our control method can effectively deal with the above problem. Figure 5 shows the results with PID controller; it can be seen that there is a large error in the position of PID control, and especially when the position reaches the maximum and minimum, the error is large.

As shown in Figure 6, our disturbance observer could observe the state information of the system with high accuracy and effectiveness. That is, the observed signals can be used in the controller design, which further illustrates the effectiveness of the method.

## 5. Conclusions

The improvement of a robust observer-based sliding mode is improved, and the efficiency of the model is improved in this paper. Aiming at the problems of high nonlinearity, strong coupling, and external interference in the system, we firstly designed a state observer for the system through the auxiliary reconstruction system, solved the state observation problem of the system, clarified the convergence condition of the observer through theoretical analysis, and verified it through simulation. Then, the position and velocity tracking problem was tackled. Considering the external disturbance, a sliding mode control of the flexible-joint manipulator was derived based on the robust observer. The control method ensures that the system state can converge exponentially to zero in finite time under different inputs and outputs. The simulation results show that the observer can quickly observe the state variables of the system. Also, combined with the sliding mode controller, the system error can quickly converge to zero. The proposed control strategy is simple and easy to implement.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References

- [1] J. Na, Y. Huang, X. Wu, G. Gao, G. Herrmann, and J. Z. Jiang, "Active adaptive estimation and control for vehicle suspensions with prescribed performance," *IEEE Transactions on Control Systems Technology*, vol. 26, no. 6, pp. 2063–2077, 2017.
- [2] W. He, X. He, M. Zou, and H. Li, "PDE model-based boundary control design for a flexible robotic manipulator with input backlash," *IEEE Transactions on Control Systems Technology*, vol. 27, no. 2, pp. 790–797, 2018.
- [3] S. Baressi Šegota, N. Anđelić, I. Lorencin, M. Saga, and Z. Car, "Path planning optimization of six-degree-of-freedom robotic manipulators using evolutionary algorithms," *International Journal of Advanced Robotic Systems*, vol. 17, no. 2, Article ID 1729881420908076, 2020.
- [4] C. Veil, D. Müller, and O. Sawodny, "Nonlinear disturbance observers for robotic continuum manipulators," *Mechatronics*, vol. 78, Article ID 102518, 2021.
- [5] S. Zhu and M. Sun, "Robust adaptive repetitive control of robotic manipulators," in *Proceedings of the IEEE 9th Data Driven Control and Learning Systems Conference (DDCLS)*, pp. 534–538, Liuzhou, China, November 2020.
- [6] M. W. Spong, "Modeling and control of elastic joint robots," *Journal of Dynamic Systems, Measurement, and Control*, vol. 109, no. 4, pp. 310–318, 1987.
- [7] Y. Yang, "A vehicle recognition algorithm based on deep convolution neural network," *Traitement du Signal*, vol. 37, no. 4, pp. 647–653, 2020.

## Retraction

# Retracted: Computational Prediction of Intrinsically Disordered Proteins Based on Protein Sequences and Convolutional Neural Networks

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] H. He and Y. Yang, "Computational Prediction of Intrinsically Disordered Proteins Based on Protein Sequences and Convolutional Neural Networks," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 4455604, 8 pages, 2021.

## Research Article

# Computational Prediction of Intrinsically Disordered Proteins Based on Protein Sequences and Convolutional Neural Networks

Hao He <sup>1</sup> and Yong Yang <sup>2</sup>

<sup>1</sup>School of Electronic and Information Engineering, Hebei University of Technology, Tianjin 300401, China

<sup>2</sup>School of Computer Science and Technology, Tiangong University, Tianjin 300387, China

Correspondence should be addressed to Yong Yang; [greatyangy@126.com](mailto:greatyangy@126.com)

Received 6 November 2021; Accepted 8 December 2021; Published 28 December 2021

Academic Editor: Daqing Gong

Copyright © 2021 Hao He and Yong Yang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Intrinsically disordered proteins (IDPs) possess at least one region that lacks a single stable structure in vivo, which makes them play an important role in a variety of biological functions. We propose a prediction method for IDPs based on convolutional neural networks (CNNs) and feature selection. The combination of sequence and evolutionary properties is used to describe the differences between disordered and ordered regions. Especially, to highlight the correlation between the target residue and adjacent residues, multiple windows are selected to preprocess the protein sequence through the selected properties. The shorter windows reflect the characteristics of the central residue, and the longer windows reflect the characteristics of the surroundings around the central residue. Moreover, to highlight the specificity of sequence and evolutionary properties, they are preprocessed, respectively. After that, the preprocessed properties are combined into feature matrices as the input of the constructed CNN. Our method is training as well as testing based on the DisProt database. The simulation results show that the proposed method can predict IDPs effectively, and the performance is competitive in comparison with IsUnstruct and ESpritz.

## 1. Introduction

Intrinsically disordered regions (IDRs) of proteins often play an important role in many biological functions while lacking a single stable structure in vivo [1]. Intrinsically disordered proteins (IDPs) can be fully or partially unstructured and generally include one or more IDRs [2]. IDPs are very common in eukaryotes. They carry out many important functions such as cell signaling and translation and can promote molecular recognition as well as protein-protein interactions [3]. Many functions of IDPs are directly associated with their structural attributes [4]. Moreover, previous studies have shown that IDPs are key players in human disease [5]. For example, 79% of cancer-related proteins are IDPs and 57% of cardiovascular disease-related proteins are IDPs [6, 7]. Besides, IDPs are also correlated with genetic diseases, neurodegenerative diseases, and Alzheimer's disease [8, 9]. Therefore, accurate prediction of IDPs is not only

important for protein description and functional annotations but also contributing to the drug design.

There are a lot of experimental techniques for identifying IDPs, such as X-ray crystallography, nuclear magnetic resonance (NMR), and circular dichroism (CD) spectroscopy. However, the experimental methods are expensive and time-consuming due to the difficulty of purification and crystallization [10]. Therefore, it is necessary to predict IDPs based on computational methods.

During the past decade, many computational methods are proposed for the prediction of IDPs. These methods can be roughly divided into three categories [11]. (1) Physicochemical-based methods: these methods are based on physicochemical properties and propensity scales of amino acids, such as FoldIndex [12], GlobPlot [13], and IsUnstruct [14]. FoldIndex predicts IDPs by calculating the ratio of average hydrophobicity to average net charge of protein sequences. GlobPlot establishes a mapping scale to reflect the relative tendency of each amino acid residue in the ordered

or disordered state based on the probability of each amino acid being in a regular secondary structure or random curl and predicts IDPs through a kernel function and filter. IsUnstruct uses the Ising model to describe the interaction between ordered and disordered states and achieves good performance. Thus, we select it as one of comparison methods. (2) Machine learning-based methods: these methods treat IDPs prediction as a binary classification problem that use positive and negative samples to distinguish ordered and disordered residues. Commonly used classification algorithms include neural network (NN), radial basis function network (RBFN), support vector machine (SVM), random forest, and so on. The representative methods in this category contain DisPSSMP [15], Dispredict [16], SPINE-D [17], ESpritz [18], RFPR-IDP [19], and so on. DisPSSMP combines RBFN and a matrix PSSMP to predict IDPs. The matrix PSSMP is a condensed position-specific scoring matrix (PSSM) according to different physico-chemical properties. Dispredict uses three kinds of features which include sequence information, evolutionary information, and structural information and predicts IDPs based on SVMs with Radial Basis Function kernel. SPINE-D is based on NN with two-hidden-layer neural network and an additional one-layer filter for smoothing prediction results. ESpritz is an ensemble of three NNs for predicting the N-terminal, internal, and C-terminal of proteins, respectively. Based on bidirectional recursive neural network, it achieves good prediction performance. Thus, it is also selected as a comparison method. (3) Meta methods: these methods combine various prediction methods into one model to further improve the prediction performance, such as MetaDisorder [20], DisCop [21], and MobiDB-lite [22]. MetaDisorder has 13 independent predictors. The final prediction result of MetaDisorder is obtained by the weighted value of the results of these 13 predictors. It possesses high prediction accuracy, but the operation is slow because it contains so many independent predictors. DisCop uses a rational design to construct a metapredictor, which selects the best performance set of 6 predictors from 20 basic predictors. The prediction results of these methods are then combined by using a regression model. MobiDB-lite is constructed based on 8 predictors, whose final consensus prediction is determined by voting.

In this paper, we propose a method to predict IDPs based on CNN and feature selection. Considering that CNN has achieved very good results in computer vision, natural language processing, and other fields, we expect to extract more hidden features by using CNN. Our previous work [23] confirms this expectation. In this paper, we improve the preprocessing process and reconstruct and train the CNN and further improve the prediction performance. The input features include sequence properties and evolutionary properties. Moreover, to highlight their specificity, sequence and evolutionary properties are preprocessed by multiple windows, respectively. Then, the preprocessed features are combined into a feature matrix as the input of the prediction model. Through preprocessing, the input information can reflect the relationships between each feature and its neighboring

features in the feature matrix and enrich the feature information extracted from protein sequences. Our prediction model contains two convolutional layers and one fully connection layer and is trained and tested on the DisProt database [24]. Finally, the proposed method is compared with two competitive prediction methods IsUnstruct and ESpritz based on the same test set.

## 2. Materials and Methods

We select 12 sequence properties and 20 pieces of evolutionary properties. The two kinds of properties are pre-processed, respectively, to highlight their specificity. Then, we train a CNN model which includes two convolutional layers and one fully connected layer to predict IDPs.

**2.1. Dataset.** The DisProt database is used to train and test the proposed methods. There are 803 protein sequences, which contain 1,254 disordered regions and 1,343 ordered regions, corresponding to 92,418 disordered residues and 315,856 ordered residues, respectively. The 803 protein sequences are randomly divided into two subsets according to the ratio of 9:1. The large dataset is the training set, containing 721 sequences with 85,184 disordered residues and 289,983 ordered residues. The small dataset is the test set, containing 82 sequences with 7,234 disordered residues and 25,873 ordered residues. Table 1 lists the specific information.

**2.2. Selected Properties.** Since the complexity of the protein sequence denotes how it can be rearranged in different ways, the low complexity regions are more likely to be disordered than ordered. We select five complexity features discussed in our previous work [25], which include topological entropy, Shannon entropy, and three amino acid propensities. Among these features, topological entropy may not be calculated directly from the protein sequence because the sequence contains 20 amino acids elements and the length of sequence does not satisfy the conditions for calculating topological entropy. Thus, before calculating the topological entropy, we map the protein sequence to 0-1 sequence. Considering the characteristics of disordered residues, large hydrophobic amino acid residues (I, L, and V) and aromatic amino acid residues (F, W, and Y) are mapped to 1, and other residues are mapped to 0. Given a protein region  $\{w(j), 1 \leq j \leq N\}$  of length  $N$ , its topological entropy  $H_T(w)$  can be calculated as follows:

$$H_T(w) = \frac{1}{N - (2^n + n - 1) + 1} \sum_{l=1}^{N-(2^n+n-1)+1} \frac{\log_2 p_{w_l^{2^n+n-1}}(n)}{n}, \quad (1)$$

where  $p_{w_l^{2^n+n-1}}(n)$  denotes the number of different subsequences with length of  $n$  in the region  $w(l) \dots w(2^n + n - 1)$ . The length of subsequences  $n$  satisfies the following:

$$2^n + n - 1 \leq |w| < 2^{n+1} + (n + 1) - 1. \quad (2)$$



TABLE 1: Datasets used in this paper.

Datasets	Disordered regions	Ordered regions	Disordered residues	Ordered residues
Training set	1,120	1,198	85,184	289,983
Test set	134	145	7,234	25,873
Total	1,254	1,343	92,418	315,856

For the same protein region  $w$ , the Shannon entropy  $H_S(w)$  can be described as follows:

$$H_S(w) = - \sum_{k=1}^{20} f_k \log_2 f_k, \quad (3)$$

where  $f_k$  ( $1 \leq k \leq 20$ ) is the probability of 20 amino acids appearing in the region  $w$ .

Three amino acid propensities are selected from GlobPlot [13], which contain Remark 465, Deleage/Roux, and Bfactor (2STD). For these three propensities, the protein region  $w$  is mapped to them, and then the average values of the mapped regions are calculated as follows:

$$\begin{aligned} m_p(w) &= [m_{p1}(w), m_{p2}(w), m_{p3}(w)] \\ &= \frac{1}{N} \sum_{l=1}^N w^{pi}(l), \quad i = 1, 2, 3. \end{aligned} \quad (4)$$

In (4), the parameter  $w^{pi}$  represents the mapped region of  $w$  with the  $i$ -th propensity, where  $i = 1, 2, 3$  correspond to Remark 465, Deleage/Roux, and Bfactor, respectively.

In addition, it has been demonstrated that disordered regions and ordered regions usually show different physicochemical properties, and thus physicochemical properties are very useful in IDPs prediction. We select seven commonly used physicochemical properties, which is collected by Jens et al [26]. They are steric parameter, polarizability, volume, hydrophobicity, isoelectric point, helix probability, and sheet probability. Following that, the average value of the mapped target region is calculated:

$$m_a(w) = [m_{a1}(w), \dots, m_{a6}(w), \dots, m_{a7}(w)] = \frac{1}{N} \sum_{l=1}^N w^{ai}(l). \quad (5)$$

Similar to (4), the parameter  $w^{ai}$  in (5) represents the mapped region of  $w$  with the  $i$ -th physicochemical properties and  $i = 1, 2, \dots, 7$ .

Finally, the PSSM is used to describe the evolutionary of each protein sequence as the evolutionary properties to enrich the information of protein sequences. They are obtained by performing three iterations of PSI-BLAST (Position-Specific Iterative Basic Local Alignment Search Tool) on NCBI (National Center for Biotechnology Information) nonredundant database with default parameters. For a protein sequence with length  $L$ , a  $L \times 20$  matrix  $M_{pssm-L}$  can be obtained. Then, the evolutionary properties of the region intercepted by the window of length  $N$  can be expressed by a  $N \times 20$  matrix  $M_{pssm-w}$ .

**2.3. Preprocessing.** In order to highlight their specificity, sequence and evolutionary properties are preprocessed, respectively. Given a protein sequence of length  $L$ , we select a window of length  $N$  and append  $N/2$  zeros to both ends of the protein sequence. For sequence properties, each region is intercepted by the window, and a 12 dimensional vector  $v_i$  ( $1 \leq i \leq L$ ) can be calculated by the following:

$$v_i = [H_T(w), H_S(w), m_p(w), m_a(w)]^T. \quad (6)$$

Assign  $v_i$  to each residue in the  $i$ -th window. Sliding the window, each residue is associated with multiple  $v_i$ . For each residue, the sequence feature vector  $x_j^{seq}$  ( $1 \leq j \leq L$ ) is the average of all  $v_i$  about it, which can be described as follows:

$$x_j^{seq} = \begin{cases} \frac{1}{j+N_0} \sum_{i=1}^{j+N_0} v_i, & 1 \leq j \leq N_0, \\ \frac{1}{N} \sum_{i=j+N_0-N+1}^{j+N_0} v_i, & N_0 < j \leq L - N_0, \\ \frac{1}{L_0 - j - N_0 + 1} \sum_{i=j+N_0-N+1}^{L_0-N+1} v_i, & L - N_0 < j \leq L. \end{cases} \quad (7)$$

In (7),  $X_j = [x_j^1 x_j^2 \dots x_j^{N_{win}} \dots x_j^{N_{win}}]$ , and  $X_j = [x_j^1 x_j^2 \dots x_j^{N_{win}} \dots x_j^{N_{win}}]$  denotes the sequence length after appending zeros.

For evolutionary properties, each region is intercepted by the window gets a  $N \times 20$  matrix  $M_{pssm-w}$ . We calculate the average value of the matrix of the intercepted region and obtain a 20 dimensional vector which is served as the evolutionary feature vector  $x_j^{evo}$  ( $1 \leq j \leq L$ ) for the central residue in the region:

$$x_j^{evo} = \left( \frac{1}{N} \sum_{l=1}^N M_{pssm-w}(l, 1:20) \right)^T. \quad (8)$$

Then, for each residue, a 32-dimensional feature vector  $x_j = [x_j^{seq}; x_j^{evo}]$  can be obtained.

In this paper, we select multiple windows to perform the preprocessing. According to the preprocessing, each residue can get a 32-dimensional feature vector  $X_j = [x_j^1 x_j^2 \dots x_j^{N_{win}} \dots x_j^{N_{win}}]$  for each window, where  $X_j = [x_j^1 x_j^2 \dots x_j^{N_{win}} \dots x_j^{N_{win}}]$  denotes the label of the window. Then, the feature vectors calculated by different windows are combined into a feature matrix. Assuming that the number



of selected windows is  $X_j = [x_j^1 x_j^2 \dots x_j^{n_{win}} \dots x_j^{N_{win}}]$ , the feature matrix of each residue  $X_j = [x_j^1 x_j^2 \dots x_j^{n_{win}} \dots x_j^{N_{win}}]$  can be expressed as follows:

$$X_j = [x_j^1 x_j^2 \dots x_j^{n_{win}} \dots x_j^{N_{win}}]. \quad (9)$$

So, each residue can obtain a feature matrix of  $32 \times N_{win}$ , where each row represents the preprocessing results of a certain feature at different windows and each column represents the preprocessing results of 32 features at a certain window. Thus, there are some correlations between the rows and columns of the feature matrix.

**2.4. Designing and Training the CNN.** We design a convolutional neural network (CNN) with two convolutional layers as well as one fully connected layer, and each convolutional layer is followed by a pooling layer, as shown in Figure 1. Since the scale of the feature matrix calculated is small, the convolution kernels are set to small scales when the CNN is designed. At the same time, because of the large number of learning samples, fewer convolution kernels and convolutional layers are selected to simplify the operation.

In the network, the activation functions of the convolutional layers are ReLu functions and the activation functions of the output layer are softmax functions. The parameters of the first convolutional layer (conv1) are set to  $3 \times 3 \times 1 \times 8$  preliminarily, where  $3 \times 3$  is the size of the convolution kernel, and 1 denotes the number of channels and 8 denotes the number of convolution kernels. Similarly, the parameter of the second convolutional layer (conv2) is  $2 \times 2 \times 8 \times 8$ . In each convolutional layer, the convolution step is 1 and performs same padding with zero. The two pooling layers use max pooling with  $2 \times 2$  filters.

In the designed CNN, the gradient descent algorithm is replaced by the Adam algorithm in the backward propagation to update parameters. In order to improve the operation speed, minibatch is used to update parameters. That is, the sample set is divided into multiple subsets of equal scale for each iteration, and each subset is used to calculate the gradient and update the parameters one by one. Finally, combined with the feature selection and extraction, Figure 2 shows the prediction procedure of the proposed method.

### 3. Results and Discussion

**3.1. Performance Evaluation.** Four metrics are used to evaluate the proposed method, which include sensitivity (Sens), specificity (Spec), weighted score (Sw), and Matthews correlation coefficient (MCC). The Sw and MCC can be computed as follows:

$$Sw = Sens + Spec,$$

$$MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (10)$$

where  $Sens = TP/(TP + FN)$ ,  $Spec = TN/(TN + FP)$ , and TP, TN, and FN as well as FP are corresponding to the

number of true positives, true negatives, false negatives, and false positives.

**3.2. Impacts of Different Number of Windows.** The lengths of windows are selected in the interval of [7, 55] firstly. The odd numbers in the interval are selected as the window lengths, which is [7, 9, . . . , 55]. Then, starting from the length of 7, we increase the step length from 1 to 4 to select windows, and thus the numbers of selected windows are 25, 13, 9, and 7, respectively. At this point, the parameters of conv1 and conv2 are set to  $3 \times 3 \times 1 \times 8$  and  $2 \times 2 \times 8 \times 8$ , respectively. The  $2 \times 2$  filters are used in the pooling layers to perform max pooling. The learning rate is 0.005. Table 1 and Figure 3 show the prediction results of 10-fold cross validation on the training set with different number of windows.

From Table 2, with the decrease in the number of windows, the value of Sens fluctuates and the value of Spec has an upward trend. And at the same time, the values of Sw and MCC increase with the decrease in the number of windows, as shown in Figure 3. Since the number of windows is inversely proportional to the distance between windows, when the window distance is small, that is, the number of windows is large, the redundancy of the calculated feature matrix is relatively high, and the prediction performance is damaged.

Considering that the prediction result of  $N_{win} = 7$  is similar to  $N_{win} = 9$ , we add some longer windows to them and make the length of the longest window around 90. In the case of  $N_{win} = 7$ , the set of window is [7, 15, . . . , 55]. We increase the number of windows by the distance between them which is equal to 8. Then, the window data set becomes [7, 15, . . . , 87] and  $N_{win} = 11$ . Similarly, in the case of  $N_{win} = 9$ , the set of window [7, 13, . . . , 55] is increased to [7, 13, . . . , 91], and the number of windows in the new set is  $N_{win} = 15$ . The prediction results of them are shown in Table 3.

From Table 3, adding some longer windows can improve the value of Sw. And the Sw of  $N_{win} = 11$  and  $N_{win} = 15$  is similar. However, the MCC of  $N_{win} = 11$  is much larger than that of  $N_{win} = 15$ , so we choose the windows in  $N_{win} = 11$ , that is, [7, 15, . . . , 87].

**3.3. Impacts of Different Number of Convolutional Layers.** Our CNN model is designed by several submodules which contain a convolutional layer and a pooling layer. Thus, when we add a convolutional layer, it is followed by a pooling layer. In this section, we add convolutional layers on the basis of the network structure in Figure 2, which includes 2 convolutional layers. For all the additional convolutional layers, the parameter is set to  $2 \times 2 \times 8 \times 8$ . Table 4 shows the prediction results of 10-fold cross validation on the training set with different convolutional layers.

In Table 4, with the increase in the number of layers, although the values of Spec fluctuated, the values of Sens show downward trend and the values of Sw and MCC also show downward trend. Therefore, we still use two convolutional layers in the prediction model.

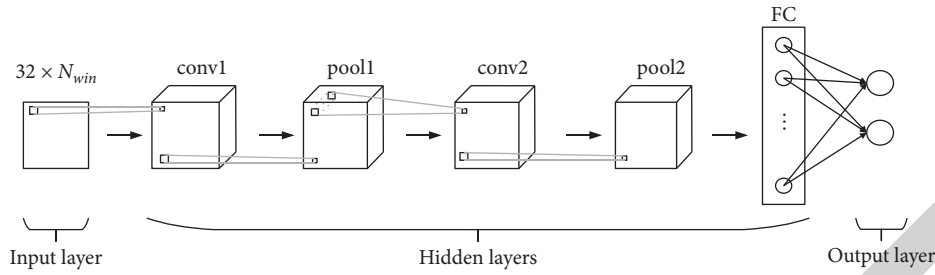


FIGURE 1: The structure of the CNN.

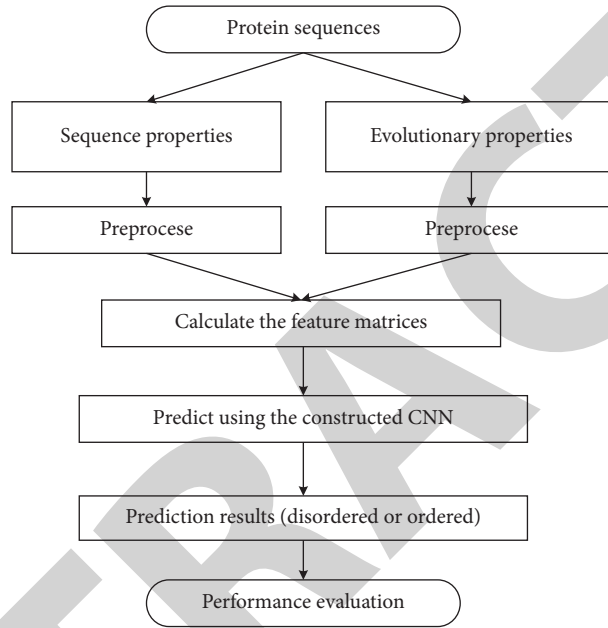


FIGURE 2: The prediction procedure of the proposed method.

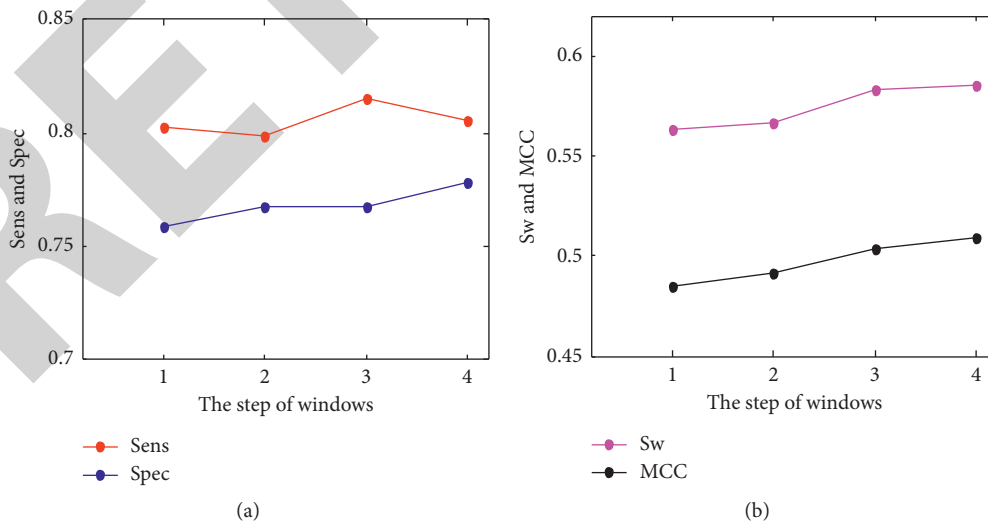


FIGURE 3: The trend of each evaluation parameter with different number of windows.

**3.4. Impacts of Different Scales of Convolution Kernels.** We change the scale of convolution kernels to analyze the influence on the prediction performance. Firstly, the number

of convolution kernel in conv1 is set to 8, and the parameter of the second convolution layer conv2 is  $2 \times 2 \times 8 \times 8$ . At the same time, the scale of the convolution kernel in conv1 is

TABLE 2: Prediction performance of different number of windows.

Step length	$N_{win}$	Sens	Spec	Sw	MCC
1	25	0.8032	0.7594	0.5626	0.4852
2	13	0.7991	0.7676	0.5667	0.4909
3	9	0.8153	0.7676	0.5829	0.5038
4	7	0.8061	0.7787	0.5848	0.5089

TABLE 3: Prediction performance of different number of windows.

Window distance	$N_{win}$	Sens	Spec	Sw	MCC
8	7	0.8061	0.7787	0.5848	0.5089
8	11	0.7833	0.8146	0.5979	0.5332
6	9	0.8153	0.7676	0.5829	0.5038
6	15	0.8914	0.7056	0.5970	0.5012

TABLE 4: Prediction performance of different convolutional layers.

Number of convolutional layers	The scale of convolutional parameter	Sens	Spec	Sw	MCC
2	$3 \times 3 \times 1 \times 8, 2 \times 2 \times 8 \times 8$	0.7833	0.8146	0.5979	0.5332
3	$3 \times 3 \times 1 \times 8, (2 \times 2 \times 8 \times 8) \times 2$	0.7813	0.7961	0.5774	0.5092
4	$3 \times 3 \times 1 \times 8, (2 \times 2 \times 8 \times 8) \times 3$	0.7654	0.7945	0.5599	0.4946
5	$3 \times 3 \times 1 \times 8, (2 \times 2 \times 8 \times 8) \times 4$	0.7513	0.8028	0.5541	0.4932

changed as shown in Table 5. Since the scale of feature matrix is only  $32 \times 11$ , the scale of convolution kernel should not be too large. We select four scales for comparison. Table 5 shows the prediction results of 10-fold cross validation on the training set with different convolution kernels in conv1. Although kernel of  $2 \times 2$  gets the highest Sens and kernel of  $4 \times 4$  gets the highest Spec, these four scales of convolution kernel obtain similar Sw and MCC. Considering that the Sw of kernel of  $2 \times 2$  is slightly higher than others, we finally set the convolution kernel of conv1 to  $2 \times 2$ .

After determining the scale of convolution kernel in conv1, we change the scale of convolution kernel in conv2. In this case, the parameter of conv1 is  $2 \times 2 \times 8 \times 8$ . Similar to the selection of conv1, we also compare the same four scales for conv2. The predicted results are shown in Table 6.

The convolution kernels of  $5 \times 5$  and  $3 \times 3$  obtain the highest Sens and Spec, respectively. However, the kernel of  $2 \times 2$  possesses much better Sw. Therefore, the scale of convolution kernel in conv2 is set to  $2 \times 2$ .

### 3.5. Impacts of Different Number of Convolution Kernels.

In this section, the number of convolution kernels is changed to analyze the influence on the prediction performance. In the design of CNN, the numbers of two convolution layers are set to be the same in analyzing the influence of the number of convolution kernels. Table 6 shows the prediction results of 10-fold cross validation on the training set when  $N_{c1} = N_{c2} = 4, 8, 16, 32$ .

From Table 7, with the increase in number of kernels, the values of Sw and MCC show downward trend, and  $N_{c1} = N_{c2} = 4$  gets better prediction performance.

TABLE 5: Prediction performance of different scales of convolution kernel in conv1.

The scale of conv1	Sens	Spec	Sw	MCC
$2 \times 2$	0.7972	0.8090	0.6062	0.5373
$3 \times 3$	0.7833	0.8146	0.5979	0.5332
$4 \times 4$	0.7553	0.8414	0.5967	0.5457
$5 \times 5$	0.7573	0.8377	0.5950	0.5423

TABLE 6: Prediction performance of different scales of convolution kernel in conv2.

The scale of conv2	Sens	Spec	Sw	MCC
$2 \times 2$	0.7972	0.8090	0.6062	0.5373
$3 \times 3$	0.7715	0.8291	0.6006	0.5422
$4 \times 4$	0.7832	0.8047	0.5879	0.5210
$5 \times 5$	0.8169	0.7734	0.5902	0.5114

TABLE 7: Prediction performance of different number of convolution kernel in conv1.

$N_{c1}, N_{c2}$	Sens	Spec	Sw	MCC
4	0.7960	0.8321	0.6281	0.5654
8	0.7972	0.8090	0.6062	0.5373
16	0.7442	0.8349	0.5791	0.5282
32	0.8226	0.7436	0.5660	0.4838

Thus, the number of convolution kernel in conv1 is fixed on 4.

Then, we only change the number of convolution kernel in conv2. Table 8 shows the prediction results when  $N_{c2} = 4, 8, 16, 32$ . From Table 8, it obtains better Sw and MCC when  $N_{c2} = 4$ . Therefore, the parameters of conv1 and conv2 are set to  $2 \times 2 \times 1 \times 4$  and  $2 \times 2 \times 4 \times 4$ , respectively.

TABLE 8: Prediction performance of different number of convolution kernel in conv1.

$N_{c1}, N_{c2}$	Sens	Spec	Sw	MCC
4	0.7960	0.8321	0.6281	0.5654
8	0.7972	0.8090	0.6062	0.5373
16	0.7442	0.8349	0.5791	0.5282
32	0.8226	0.7436	0.5660	0.4838

TABLE 9: Prediction performance of different number of convolution kernel in conv1.

Methods	Sens	Spec	Sw	MCC
Our method	0.7264	0.8301	0.5565	0.5060
IsUnstruct	0.7513	0.7855	0.5368	0.4711
ESpritz	0.7255	0.8135	0.5389	0.4840

3.6. *Comparison with Other Methods.* Our method is compared with other two state of the art methods IsUnstruct and ESpritz in this section. Table 9 shows the prediction performance of three methods based on the test set. The prediction results of IsUnstruct and ESpritz are obtained by their respective network predictors. As shown in Table 9, our method gets the best Spec and similar Sens as ESpritz and thus obtains higher Sw and MCC.

## 4. Conclusions

In this paper, we propose a method to predict IDPs based on CNN and feature selection. Sequence properties and evolutionary properties are used to describe the differences between disordered residues and ordered residues. To highlight their specificity, sequence and evolutionary properties are pre-processed by 11 windows from length 7 to length 87, respectively. Then, the preprocessed features are combined into a feature matrix as the input of the prediction model. CNN can reflect the relationship between each feature and its neighboring features in the protein feature matrix and find out more information from different features and thus enrich the information proposed by protein sequences. Thus, we construct a CNN prediction model with two convolution layers and one fully connected layer, and each convolution layer is followed by a pooling layer. The parameters of each convolution kernel are set to  $2 \times 2 \times 1 \times 4$  and  $2 \times 2 \times 4 \times 4$ , respectively. The simulation results show that the prediction performance of the proposed method gets better Sw and MCC than two competitive prediction methods IsUnstruct and ESpritz.

## Data Availability

The datasets supporting the conclusions of this article are available on the DisProt database [24] (<http://www.disprot.org/>).

## Additional Points

In this paper, the authors add the technical details and experiments, improve the preprocessing process, retrain the prediction network, and improve the prediction performance.

## Disclosure

The previous work “Prediction of Intrinsically Disordered Proteins with Convolutional Neural Networks Based on Feature Selection” [23] is published in 2021 International Conference on Computer Engineering and Artificial Intelligence. The funding bodies have no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported by the Hebei Province University Science and Technology Research Project (No. QN2021038).

## References

- [1] C. J. Oldfield, V. N. Uversky, and L. Kurgan, “Predicting functions of disordered proteins with MoRFPred,” *Methods in Molecular Biology*, vol. 1851, pp. 337–352, 2019.
- [2] Z. Peng, Q. Xing, and L. Kurgan, “APOD: accurate sequence-based predictor of disordered flexible linkers,” *Bioinformatics*, vol. 36, no. Suppl\_2, pp. i754–i761, 2020.
- [3] C. C. Hsu, M. J. Buehler, and A. Tarakanova, “The order-disorder continuum: linking predictions of protein structure and disorder through molecular simulation,” *Scientific Reports*, vol. 10, no. 1, p. 2068, 2020.
- [4] R. Van der Lee, M. Buljan, B. Lang et al., “Classification of intrinsically disordered regions and proteins,” *Chemical Reviews*, vol. 114, no. 13, pp. 6589–6631, 2014.
- [5] A. F. Faustino, G. M. Barbosa, M. Silva et al., “Fast NMR method to probe solvent accessibility and disordered regions in proteins,” *Scientific Reports*, vol. 9, no. 1, p. 1647, 2019.
- [6] A. K. Dunker, S. E. Bondos, F. Huang, and C. J. Oldfield, “Intrinsically disordered proteins and multicellular organisms,” *Seminars in Cell & Developmental Biology*, vol. 37, pp. 44–55, 2015.
- [7] I. Staneva, Y. Huang, Z. Liu, and S. Wallin, “Binding of two intrinsically disordered peptides to a multi-specific protein: a combined Monte Carlo and molecular dynamics study,” *PLoS Computational Biology*, vol. 8, no. 9, Article ID e1002682, 2012.
- [8] U. Midic, C. J. Oldfield, A. K. Dunker, Z. Obradovic, and V. N. Uversky, “Protein disorder in the human diseaseome: unfoldomics of human genetic diseases,” *BMC Genomics*, vol. 10, no. S1, p. S12, 2009.
- [9] V. N. Uversky, C. J. Oldfield, and A. K. Dunker, “Intrinsically disordered proteins in human diseases: introducing the D2 concept,” *Annual Review of Biophysics*, vol. 37, no. 1, pp. 215–246, 2008.
- [10] C. J. Oldfield, E. L. Ulrich, Y. Cheng, A. K. Dunker, and J. L. Markley, “Addressing the intrinsic disorder bottleneck in structural proteomics,” *Proteins: Structure, Function, and Bioinformatics*, vol. 59, no. 3, pp. 444–453, 2005.
- [11] Y. Liu, “A comprehensive review and comparison of existing computational methods for intrinsically disordered protein and region prediction,” *Briefings in Bioinformatics*, pp. 1–17, 2017.

## Research Article

# Modeling the Use of Online Knowledge Community: A Perspective of Needs-Affordances-Features

Zeyu Jiao <sup>1</sup>, Jianbin Chen <sup>2</sup>, and Eunjin Kim <sup>1</sup>

<sup>1</sup>Kyonggi University, Suwon 443-760, Gyeonggi-do, Republic of Korea

<sup>2</sup>Beijing Union University, Beijing 100101, China

Correspondence should be addressed to Jianbin Chen; [jianbin.chen@buu.edu.cn](mailto:jianbin.chen@buu.edu.cn)

Received 26 October 2021; Accepted 16 November 2021; Published 27 December 2021

Academic Editor: Daqing Gong

Copyright © 2021 Zeyu Jiao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the support of network information technology, the Online Knowledge Community (OKC) has emerged. Among different OKC applications, some entered into the new era of popular knowledge production, while others experienced the process to decline. In order to solve the dilemma faced by the OKC platforms, the needs-affordances-features (NAF) perspective on OKC is proposed by taking Zhihu, one of the most popular OKC applications in China as an example. According to NAF, the psychological needs of individuals motivate the use of Zhihu to the extent to which Zhihu offers affordances that satisfy the individuals' needs. By collecting data through questionnaires and using SPSS and AMOS for data analysis, the relationship between individuals' psychological needs and Zhihu affordances is identified. This paper generates two aspects of implications. In terms of theoretical implications, a more comprehensive framework is developed for the affordances of OKC as a whole, and the NAF model is leveraged to identify related psychological needs which motivate the use of a specific OKC application—Zhihu. Further research can leverage NAF to identify different OKC platform features which satisfy the psychological needs of their targeting users to optimize the system of OKC platforms. As for practical implications, by building the relationship between the affordances of OKC platforms and users' psychological needs, marketers have to realize that although the digital platform system has a certain degree of imitability, the value positioning, user community, and core capabilities behind those platforms are all different. Thus, the platform system must also be differentiated. In order to determine the appropriate business model, a clear understanding of these organizational features should be identified.

## 1. Introduction

Knowledge management has attracted the attention of many researchers due to its close relationship with organizational performance [1]. In the Internet age, the “Online Knowledge Community” (OKC) that integrates the functions of “knowledge sharing” and “online social networking” has emerged. With the opening of registration restrictions, a large number of ordinary users have poured in. Representative OKC platforms such as Douban and Zhihu have entered the second stage of development from the early stage of highly acclaimed and rapid development: the era of popular knowledge production [2]. The vast information capacity and random insertion and editing at any time have significantly increased the amount of information and the

level of confusion, which changed the traditional online communication methods [3]. At the same time, the operation of knowledge communities is facing a complex situation. For example, although Wukong Questions and Answers (Q&A) application persuaded over about 300 influencers of Zhihu to join with a high salary in 2017, the server of Wukong Q&A ceased operation in 2021 [4]. During the same period, Zhihu ushered in its tenth birthday celebration, and there were rumors of listing on the stock market. There is the success of Zhihu, while there is also the shutdown of Encarta, the failure of Google Answers, the decline of Wikipedia [5, 6], and the quit of Wukong Q&A, which operated less than four years independently.

Of these OKC platforms, why do high-quality creators prefer Zhihu? From a social-technical perspective, the OKC



platform is a collection of users' behavioral possibilities and needs in social media and organizational environments [7]. Its value is created by digital technology and is cocreated by the interaction between the user, the technology, and the purpose of use [8]. Ongus [9] also mentioned that the quality and relevance of technology resources and information content are closely related to users' needs. Therefore, the platform functions or resources enabled by platform features are essential for users' needs. However, the current literature has not yet studied this mechanism, resulting in a "black box" of the platform system. In the research of Facebook using motivations, the needs-affordances-features (NAF) model was first suggested by focusing on platform characteristics in the context of social media [10]. According to NAF, the psychological needs of individuals motivate the use of applications to the extent to which these applications offer affordances that satisfy individuals' needs (ibid). As OKC is similar to social media given the "online social networking" functions, the NAF model is proposed to address the following questions:

- (1) What innate psychological needs do people seek to satisfy through OKC?
- (2) What are the main affordances provided by OKC?
- (3) Which of these OKC affordances meet which psychological needs?

In this paper, according to the literature review, the needs-affordances-features (NAF) model is used to identify the relationship between psychological needs, affordances, and features. Based on the research of psychological needs, OKC affordances, and Zhihu features, hypotheses of this paper are suggested, and the conceptual framework is sorted out. In the final, from data collection and analysis, the relationship between individuals' psychological needs and Zhihu affordances is identified. The result shows that each of the Zhihu affordances is likely related to fulfilling some psychological needs. For example, the need for autonomy can motivate people to use Zhihu that has the affordance of browsing others' content, while the need for relatedness can motivate the use of Zhihu that has affordances of self-presentation, relationship formation, and meta-voicing.

In terms of the research gap, although current research has noticed that the platform's community ecology and value cocreation may play an essential role in optimizing the platform, the platform is mainly regarded as an intermediary connection point and ecological core node [11, 12]. That is to say, there is a lack of analysis of the internal micro-mechanism based on the platform's individualization and differentiation. Therefore, this paper will address the "black box" of the platform system. With the consideration of platform affordances, the interaction relationship between the platform and users will be studied from the dimensions of core functions and resources to enrich the theoretical system of the platform further and achieve theoretical innovation. In the view of practice, from the fact that Wukong Q&A has invested billions in the construction of the community, although it still fell to the ground, it can be seen that different platform affordances could create different

community ecology. The research of this topic can guide more platforms to design more effective OKC applications under the premise of unifying user needs and organizational characteristics, which could lead to avoiding waste of social resources and improving the success rate of Internet innovation.

## 2. Literature Review

*2.1. Online Knowledge Community (OKC).* Online Knowledge Community (OKC) refers to a virtual space where users scattered in different regions carry out knowledge exchange and interpersonal interaction with the support of network information technology [13], which is a typical social-technical system [14]. OKC has the collaborative ordering of the social and knowledge systems, and users are the driving force of ordering the two systems [3, 15].

Currently, OKC research focuses on three hot areas: user research, content research, and community research [2]. User research mainly explores OKC users' behavior characteristics and patterns and analyzes user motivation for participation and the interaction between users. Content research focuses on content quality measurement and evaluation, semantic analysis, and topic recognition, influencing factor analysis, relationship research, and so forth. Community research is about conducting research on community operation mechanisms, content generation models, and community comparisons. The representative viewpoints are as follows: the field of information management emphasizes the impact of IT technology on knowledge proliferation [16]; the field of sociology pays more attention to the impact of trust on online interaction [17]. Guan et al. [18] believed that knowledge inflow is the key and examine the antecedents of knowledge contribution behavior in multiple dimensions. In contrast, Lai et al. [19] and Zhao and Huang [20] believed that knowledge search is equally important. They examine the antecedents of knowledge search behavior from the perspectives of behavior attitude, subjective norms, and ability beliefs. In terms of user research, opinion leaders or leading users usually have more experience or mastered more professional knowledge, so their behaviors are more exploratory or creative, and they are easier to gain trust. In addition, knowledge contribution is their important active behavior [21].

*2.2. Affordance Theory.* Affordance originally refers to the support that objective things can provide for a certain behavior, that is, action possibilities permitted by properties of objects [22]. In recent years, it has become more popular in organizational research and can be used to better understand the impact of the combination of new technologies and organizational characteristics on organizational innovation and operations [23]. "Affordance" provides not only a strong theoretical perspective for studying the relationship between technology and personnel in an organization but also a better language for the structured and patterned description of specific practices [24].



As a commonly used communication method, social media affordances have an important influence on the organization's communication process, employee and user behavior, and psychology, so it is currently a hot topic of affordance research. Postigo [25] first analyzed, from the perspective of social-technical interaction, how YouTube guides users to conduct behaviors that are beneficial to the commercial interests of the platform through the design of platform architecture. Rice et al. [7] defined social media affordances as the relationship between the action possibility and the need (or purpose) perceived by users aggregated in social media and the organizational environment under the constraints of the potential features/functions of the social media platforms. Majchrzak and Faraj [23] studied the relationship between the different ways through which employees participate in publicly visible conversations on social media platforms and knowledge sharing. Sun et al. [26] divided the use of social media platforms into three categories: interactivity (conversations between users and comments on related information), accessibility (users can obtain relevant information), and sociality (users establish connections with other users).

There are many research results around the dimension of affordances. Treem and Leonardi [27] proposed four dimensions of social media affordances, visibility, associating, editability, and persistence, which have been commonly used [24, 26, 28]. Majchrzak and Faraj [23] proposed four dimensions of affordances for knowledge sharing (meta voicing, trigger attending, network-informed associating, generative role-taking). In addition, there are six dimensions of functional media affordances [7], ten dimensions of communication affordances [29], and six dimensions of social business technology affordances [30]. As for new media affordances, Pan and Liu [31] proposed production affordances (including editability, reviewability, replicability, scalability, and associability), social affordances (including greet-ability, emotion-ability, coordinate-ability, and connect-ability), and mobile affordances (including portability, availability, locatability, and multimodality), which reflect the initiative of the information prosumer.

The most integrated literature is proposed by Karahanna et al. [10], who divide social media affordances into ego-centric and allocentric affordances based on the theory of self-determination and psychological ownership. In addition, the research framework for social media adoption and use is proposed, which is the needs-affordances-features (NAF) model, and an empirical study of affordances is conducted with Facebook as an example. Zhang and Huang [32] discussed the realization of platform affordances from aspects of technology affordances and social affordances and emphasized that users' perceptions and actions must be considered when gaining insight into the complex relationship between technology and society.

Based on the above literature, OKC affordances could be divided into platform affordances and product affordances in this research. Platform affordances include technology affordances and social affordances. However, this paper focuses on the relationship between individuals' psychological motivations and OKC affordances. Compared to

technology affordances which provide technical support for platform affordances, social affordances can better reflect the relationship between platform affordances and people's psychological needs for using a certain OKC platform. As a result, social affordances will be used as platform affordances in this research, which is realized by platform functions, such as Zhihu platform features. Product affordances focus on the affordances of OKC knowledge resources, referred to as knowledge affordances (see Table 1).

*2.3. NAF Perspective for OKC.* The psychology literature shows that people are driven to engage in activities that fulfill their innate psychological needs. Therefore these psychological needs are stimulating states that act as motivations for action [33]. Based on this premise, a needs-affordances-features (NAF) perspective shows that innate psychological needs motivate people to use OKC apps, which have affordances to satisfy their psychological needs potentially. For example, OKC apps provide the affordance to connect with others, through features such as "Following" and "Chatting" on Zhihu. This affordance can be used to satisfy people's psychological needs for relatedness. Therefore, this psychological need drives them to participate in using Zhihu features that provide this affordance to meet this psychological need (see Figure 1).

*2.4. Relationship between Psychological Needs and OKC Affordances.* Although there is no research to categorize psychological needs in the OKC context, research on human motivations shows that "motivation can be conceived as a duality" [34]. On the one hand, people focus on the self. On the other hand, they strive for relationships between the self and others. Hermans [35] suggested the former motivation as the S-motive and the latter as the O-motive. Based on self-determination theory (SDT) [33, 36] and psychological ownership theory (POT) [37–39], seven psychological needs are listed, which span the two polarities and illustrate people's self-focused and other-focused motivations. Thus, these two theories are synthesized in this paper to show psychological needs in a more comprehensive way. According to the terminology of Hermans [35], these seven psychological needs are divided into the self-focused group, also called S-needs (need for autonomy, competence, having a place, coming to know the self, maintaining continuity of self-identity, and one aspect of expressing self-identity), and the other-focused group, also called O-needs (need for relatedness and one aspect of expressing self-identity).

According to Karahanna et al. [10], there are 12 social media affordances, which constitute egocentric affordances (self-presentation, content sharing, and interactivity) and allocentric affordances (presence signaling, relationship formation, group management, browsing other's content, meta-voicing, communication, collaboration, competition, and sourcing). As OKC is a typical social-technical system equipped with the function of "online social networking," these 12 social media affordances will be used to analyze the OKC social affordances. However, in addition to social

TABLE 1: OKC affordances concept system.

	Affordance concept system		Platform factors
OKC affordances	Platform affordances	Social affordances	Platform functions
	Product affordances	Knowledge affordances	Platform resources

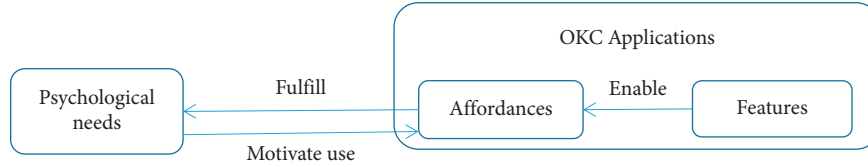


FIGURE 1: Logic of the NAF perspective.

affordances, OKC has its specific product, which is knowledge. According to Shi et al. [40], knowledge affordances are defined as unique attributes of knowledge that meet the needs of consumers, which can be measured in 4 dimensions (reliability, economies, selectivity, and uniqueness). In order to map the OKC affordances to psychological needs, Table 2 is shown based on former research [10] and six OKC researchers (see Table 3 for details of term definitions).

**2.5. Literature Gaps.** First, there is a lack of platform features in the context of OKC research. At present, the OKC research has three hot areas: users, content, and communities/platforms, and OKC affordances have also attracted much attention. However, the OKC research focuses on the field of subrole characteristics and its socialization mechanism and lacks attention to platform differences. Second, people’s attention to the role of psychological needs in motivating the use of the OKC platforms is limited. In particular, there has not been a comprehensive theoretical-based attempt to determine a set of psychological needs that are salient in the OKC environment. In most studies, the focus is not on theorizing around psychological needs, but the needs variable is one of the many variables included in the research model. Third, current studies did not adopt a systematic method to identify the OKC affordances that meet these needs. On the contrary, most people measure need satisfaction to confirm whether affordances meet people’s psychological needs. The focus on the relationship between psychological needs and OKC affordances provides valuable operational design guidelines for information system researchers and practitioners. Our research aims to address these gaps.

### 3. Research Model and Hypotheses

There are three steps of developing the NAF model in the use of the Zhihu application. Firstly, to understand what psychological motivations of using this app, the affordances provided by the Zhihu platform should be identified. In 2011, the mission of Zhihu was officially launched, which is “allowing people to share knowledge, experience and insights better and find their own answers” [58]. After nearly ten years of development, Zhihu announced the brand

renewal and upgrade at the 2021 New Knowledge Youth Conference. The brand slogan has been upgraded from “If there is a question, go to Zhihu” to “If there is a question, there will be an answer” (ibid). These show that self-presentation, content sharing, browsing others’ content, and sourcing are salient affordances of Zhihu. Based on the functions of chatting and live Q&A, liking or collecting others’ content, and voting for what others posted, it is obvious that communication and meta-voicing are also affordances of Zhihu. In addition, there is a function of “Quanzi” in Zhihu, which offers a community for the group of people who have common interests, where they could share and communicate about their experience and opinions in their own circle. As a result, relationship formation is also one of the social affordances of Zhihu. After identifying salient affordances, the most popular features of Zhihu were selected from a synthesis of research on the condition that affordances are provided by specific features of an object [52, 57, 59, 60]. With the combination of its knowledge properties, six researchers with research expertise in OKC (three faculty and three doctoral students) are asked to map the features to the 13 OKC affordances listed in Table 2. Table 4 illustrates the result of the mapping of Zhihu features to Zhihu affordances. Therefore, Zhihu affordances are identified as self-presentation, content sharing, relationship formation, browsing others’ content, meta-voicing, communication, sourcing, and knowledge attributes.

In order to predict which Zhihu affordances meet which psychological needs, the hypotheses are suggested as follows.

According to Deci [61], autonomy is defined as people’s innate psychological needs to act authentically in their own life. Everyone is inherently eager for autonomy [62]. They participate in activities, not because of social norms or pressures but because they are free to choose [33, 63]. Therefore, we suggest that self-presentation, content sharing, relationship formation, browsing others’ content, and sourcing are affordances that motivate users to use Zhihu. For example, those affordances allow individuals of Zhihu to choose how to present themselves (through uploading video and photos, disclosing locations), what content to share, whom to follow, what kind of community to join, and what questions to suggest or answer, which leads to hypothesis 1.

TABLE 2: Mapping of the psychological needs to OKC affordances.

OKC Affordances		Needs						
		S-NEEDS			O-NEEDS			
		A	C	HP	CK	MC	ES	R
Social Affordances	Self-presentation	√		√		√	√	√
	Content Sharing	√		√		√	√	
	Interactivity	√		√				
	Presence Signaling							√
	Relationship Formation	√					√	√
	Group Management		√					√
	Browsing Others' Content	√			√			√
	Meta-voicing		√		√		√	√
	Communication						√	√
	Collaboration		√		√			√
	Competition		√		√			
	Sourcing	√	√					
Knowledge Affordances	Knowledge attributes		√					

Note: A = autonomy; C = competence; HP = having a place; CK = coming to know the self; MC = maintaining continuity of self-identity; ES = expressing self-identity; R = relatedness.

H1: The need for autonomy can motivate the use of Zhihu that has these affordances: self-presentation, content sharing, relationship formation, browsing others' content, and sourcing.

In 1991, Deci and Ryan illustrated that relatedness is individuals' innate psychological need to be connected to others. In 2000, Deci and Ryan extended the definition as the need "to love and care and to be loved and cared for." Thus, we posit that a set of Zhihu affordances—self-presentation, relationship formation, browsing others' content, meta-voicing, and communication—can help people satisfy the need for relatedness by creating social connections with others. This can be realized by reaching a lot of users, joining an online group, knowing what others are doing, reacting to others' posts, and so forth [64, 65]. For example, in the context of Zhihu, users can increase social interactions by posting personal information (self-presentation), following a user (relationship formation), browsing others' posts (browsing others' content), voting for a post (meta-voicing), sending an instant message to others (communication), and so forth. Therefore, the H2 is proposed as follows.

H2: The need for relatedness can motivate the use of Zhihu that has these affordances: self-presentation,

relationship formation, browsing others' content, meta-voicing, and communication.

Bauer and McAdams [66] analyzed that competence refers to people's innate psychological needs to deal with their environment effectively. It is the psychological needs that can be achieved by having an impact on individuals or the environment. White [67] showed that individuals need to feel competent in controlling or altering the environment and finding opportunities to increase their own knowledge or skills. Therefore, people who have a high demand for competence are more likely to seek affordances that provide them with opportunities to apply or expand their knowledge in the environment. Thus, we suggest that the affordances of meta-voicing, sourcing, and knowledge attributes can fulfill individuals' needs for competence by providing feedback to others' posts, responding to others' questions, searching answers, subscribing to valuable content, attending online courses, and so on. Thus, hypothesis 3 is suggested.

H3: The need for competence can motivate the use of Zhihu that has the following affordances: meta-voicing, sourcing, and knowledge attributes.

In terms of having a place, it refers to people's innate psychological needs to possess a place where they can create their own space [68, 69]. Pierce et al. [70] suggest

TABLE 3: OKC affordances.

Affordances	Definitions	Example features
	Social affordances	
Self-presentation	Users can display and present information related to themselves. This includes sharing information that somehow portrays users and shows who they are, their values and preferences, their expertise, etc. Updating descriptive information about themselves, such as gender, occupation, and location; and posting content involving pictures and videos related to themselves [41–44]	Posting my own content on Zhihu; sharing my own video on YouTube; updating my profile on Zhihu; writing personal opinions on Zhihu
Content sharing	Users can share and distribute content unrelated to them to others (e.g., sharing posts, video) [27, 45, 46]	Sharing links of other people’s articles on Zhihu; sharing others’ videos and photos on Instagram
Interactivity	Users can walk around in real time and change their virtual environment (e.g., to build in-world artifacts) [41]	Moving around in World of Warcraft
Presence signaling	Users can indicate their existence and know whether other users are accessible [43, 44, 47]	“Who is online” on Instagram
Relationship formation	Users can establish relationships with others, including joining groups or online communities [27, 45]	Following other users on Zhihu; joining an online community (e.g., “Quanzi” on Zhihu)
Group management	Users can form groups and online communities and manage them. The focus is on the management or administration of groups [45, 46]	Illegal content management on Zhihu, forming a “Quanzi” on Zhihu
Browsing Other’s content	Users can view others’ content and receive alerts to pay attention to others’ content [41, 42, 45]	Browsing others’ content on Zhihu, receiving notifications on LinkedIn
Meta-voicing	Users can participate in online conversations by responding to other people’s status, profile, content, and activities online and viewing other people’s responses to their status, profile, content, and activities. In meta-voice, the user “not only has to express his or her opinion, but also add meta-knowledge to content already online.” [23, 48]	Voting on a post on Zhihu, answering questions on Zhihu, liking what others post on Zhihu
Communication	Users can chat or send messages with others directly [41, 44, 47]	Chatting on Zhihu, communicating with others on Google+
Collaboration	Users can collaborate with others, such as collaborating with others to create content on Wikipedia [23, 43, 48]	Adding, deleting, and editing content on Zhihu
Competition	Users are able to compete with others, which includes competing in online games [10]	Completing tasks in World of Warcraft
Sourcing	Users are able to ask for resources or funds, including meeting others’ requests for funds or resources [10]	Asking or answering questions on Zhihu
	Knowledge affordances	
Knowledge attributes	Reliability: It refers to the extent to which the answers on social Q&A websites make users feel trustworthy and reliable [49]. Users think that the answer is of high quality only when they believe that the source and content of the answer are reliable [50]	The reliability of Zhihu content is reflected in its questions, answers, articles, videos, pictures, etc.
	Selectivity: Users can subscribe to specific content or sources of information [51]	Zhihu involves popular and unpopular content in multiple sections, and the content knowledge within each section is highly subdivided [52]
	Economies: It means that the subject obtains relatively maximum benefits with relatively minimum investment to obtain benefits most economically and meet the needs of survival and development [53]	Zhihu provides users with a free Q&A community [54]. Users can spend less money to ask questions to experts in related fields
	Uniqueness: It is defined as individuals pursuing unique characteristics different from others by acquiring, using, and disposing of consumer goods [55]. Novelty is a concept closely related to uniqueness. Novelty refers to the extent to which the answers on social Q&A websites make users feel innovative. Innovative answers will bring new ideas to users and will also be regarded as high-quality answers by users [49, 56]	In-depth content production is different from the knowledge provided by other Q&A platforms. Online and offline knowledge products are carried out at the same time [57]



TABLE 4: Mapping of Zhihu features to Zhihu affordances.

Zhihu platform features	1	2	3	4	5	6	7	8
Uploading own content	√							
Sharing other's content		√						
Watching live				√				
Commenting on other's post					√			
Asking or answering questions							√	
Liking or collecting what others posted					√			
Voting for what others posted					√			
Chatting							√	
Joining an online community			√					
Browsing other's content				√				
Following other users			√					
Searching answers							√	
Writing the column		√						
Zhihu product features								
Reliability								√
Selectivity								√
Economies								√
Uniqueness								√

1-7: self-presentation, content sharing, relationship formation, browsing others' content, meta-voicing, communication, sourcing, and knowledge attributes.

that a sense of having a place can be partly realized by the personalization of individuals' surroundings. Applying this definition to cyberspace, users can fulfill their needs for having a place by investing time, energy, or emotion in creating their own virtual world through self-presentation (such as posting personal photos or videos on Zhihu) and content sharing (such as sharing others' content on Zhihu). In this way, a sense of having a place can be created by engaging in personalizing the cyberspace. As a result, hypothesis 4 can be proposed.

H4: The need for having a place can motivate the use of Zhihu that has these affordances: self-presentation and content sharing.

As for coming to know the self, Pierce et al. [71] illustrated it as individuals' innate psychological needs to identify who they are and learn about themselves. Self-identity can be developed through self-awareness, such as comparing themselves with others when people interact with their surroundings [72]. In addition, the sense of coming to know the self can be achieved by receiving feedback from others and seeing how other people think of themselves [73]. Thus, people can reflect their feelings, thoughts, and behaviors, which enable them to discover what kind of people they are. Therefore, the psychological needs for coming to know the self can be fulfilled by these affordances: browsing others' content (which helps individuals compare themselves to others) and meta-voicing (which enables individuals to see reflected appraisal). Therefore, hypothesis 5 is suggested.

H5: The need for coming to know the self can motivate the use of Zhihu that has these affordances: browsing others' content and meta-voicing.

Expressing self-identity is defined as people's innate psychological needs to communicate their identities with others [71]. For example, Zhihu users can satisfy their needs for expressing self-identity by self-presentation (through disclosing their personal information such as profile photos and education), content sharing (through sharing articles), relationship formation (through following other Zhihu users or joining online communities), meta-voicing (through commenting on or voting for others' posts), and communication (through chatting directly with other Zhihu users). Therefore, hypothesis 6 can be proposed.

H6: The need for expressing self-identity can motivate the use of Zhihu that has these affordances: self-presentation, content sharing, relationship formation, meta-voicing, and communication.

The final psychological needs suggested in this paper are maintaining continuity of self-identity, which shows individuals' innate psychological needs to maintain emotional connections between past and present [71]. Taking Zhihu as an example, self-presenting and content sharing enable users to see who they were and who they are by the questions they asked or answered, articles they shared, photos and videos they posted, and so forth. All the content constitutes one's self-identity. Thus, we posit hypothesis 7.

H7: The need for maintaining continuity of self-identity can motivate the use of Zhihu that has these affordances: self-presentation and content sharing.

Based on the above hypotheses, the conceptual framework can be shown in Figure 2.

## 4. Research Methods

**4.1. Questionnaire Design.** Based on previous research of psychological needs and Zhihu features, questionnaire 1 and questionnaire 2 are designed to test the model of this paper. In order to measure seven variables of psychological needs, 20 question items are suggested. A adopts three-question items from Deci [61] and Deci and Ryan [33, 62, 63]. R is measured by the scale of the research from Deci and Ryan [63, 74], Jenkins-Guarnieri et al. [64], and Seder & Oishi [65]. C adopts the scale from Bauer and McAdams [66] and White [67]. HP adopts two-question items from Barki et al. [68], Harrison and Barthel [69], and Pierce et al. [70]. CK is measured by the scale of Pierce et al. [71], Festinger [72], and Mead [73] while ES and MC adopt the source from Pierce et al. [71] to be the measurement scale (see in questionnaire 1). In questionnaire 2, 17 question items are suggested to measure the extent to which the use of or feeling Zhihu features. In this study, every measurement item is measured using the five-point Likert scale. 1 for "strongly disagree," 2 for "disagree," 3 for "undecided," 4 for "agree," and 5 for "strongly agree."

TABLE 5: Scale items.

Construct	Abbr.	Questionnaire 1 Items	Sources
Autonomy	A	Psychological needs I need freely choose what I want. I need present myself in my own way	Deci [61]; Deci and Ryan [33, 62, 63]
Relatedness	R	I have the need to act freely I need participate in group activities I need connect with other people	Deci and Ryan [63, 74]; Jenkins-Jenkins-Guarnieri et al. [64]; Seder and Oishi [65]
Competence	C	I need be close to many people I need expand my knowledge I need show my capabilities I need feel competent	Bauer and McAdams [66]; White [67]
Having a place	HP	I need a place of my own I need create a place which makes me feel like home	Barki et al. [68]; Harrison and Barthel [69]; Pierce et al. [70]
Coming to know the self	CK	I need to know who I am. I need to know how other people think of me.	Pierce et al. [71]; Festinger [72]; Mead [73]
Expressing self-identity	ES	I need to develop a sense of self-identity. I need to show people who I am. I need to express my personality.	Pierce et al. [71]
Maintaining continuity of self-identity	MC	I need to show my self-identity. I need to compare my past and today. I need to show other people who I were and who I am. I need my past to be an important part of my self-identity.	Pierce et al. [71]

All items were measured on a 5-point scale: 1 = strongly disagree; 5 = strongly agree



TABLE 5: Continued.

Questionnaire 2	
<i>Zhihu features</i>	
The extent to which the use of or feeling the following Zhihu features (all items were measured on a 5-point scale Zhihu platform features: 1 = never, 5 = very often Zhihu product features: 1 = strongly disagree, 5 = strongly agree)	
	F1 Uploading own content
	F2 Sharing other's content
	F3 Watching live
	F4 Commenting on other's post
	F5 Asking or answering questions
	F6 Liking or collecting what others posted
	F7 Voting for what others posted
	F8 Chatting
	F9 Joining an online community
	F10 Browsing other's content
	F11 Following other users
	F12 Searching answers
	F13 Writing the column
	F14 Reliability
	F15 Selectivity
	F16 Economies
	F17 Uniqueness

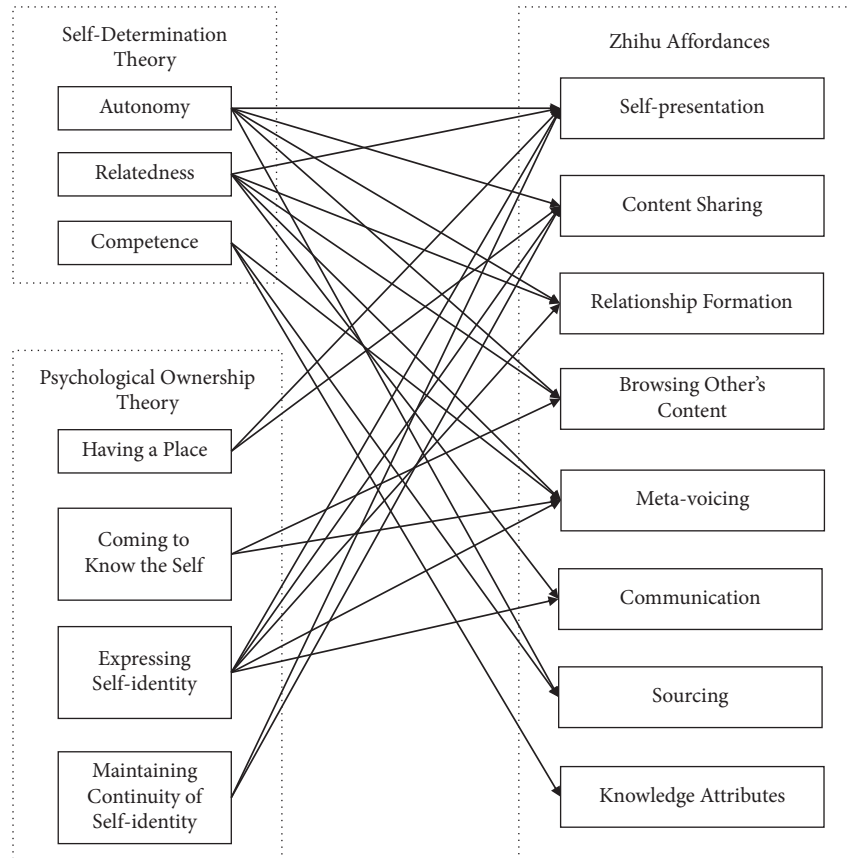


FIGURE 2: NAF model for Zhihu affordances use.

4.2. Data Collection. Taking into account the time constraints and sample size, the research scope is mainly selected in the first-tier cities in China that have a large number of people using Zhihu. First, the questionnaire was distributed and collected in a small range through the social network of teachers and classmates to conduct a preliminary

test of the questionnaire. After the questionnaire was revised, a professional survey company was commissioned to invite Zhihu users in China's first-tier cities to respond online. The survey period was from June 2021 to July 2021. A total of 300 questionnaires were returned. After excluding uncompleted and regular answers, 300 questionnaires were

finally returned, with 208 valid questionnaires, and the effective response rate was 69.3%. Based on the final sample, this paper mainly adopts SPSS26.0 and Amos24.0 software to analyze the sample data as shown in Table 5.

## 5. Results and Analysis

In this study, SPSS26.0 software and Amos24.0 software are used to analyze the research samples. Among them, the analysis methods involved in SPSS26.0 include reliability test and regression analysis, while Amos24.0 is used for confirmatory factor analysis.

*5.1. Reliability and Validity Test.* Reliability refers to the consistency or stability of the measurement results obtained by measurement tools [75]. In this paper, Cronbach's Alpha reliability test is used for testing variables. In terms of validity, it refers to the degree to which measurement tools can accurately measure the things that need to be measured [76]. This paper uses confirmatory factor analysis to test the validity of variables, which is mainly used to verify convergent validity and discriminant validity. According to Hair [76], the absolute value of factor loading should be at least 0.5 or more, and the best index value should be more than 0.7. In addition, the average variance extraction (AVE) index value should be more than 0.5. The value of the construct reliability should be higher than 0.7 to judge whether it has convergent validity. Fornell and Larcker [77] pointed out that the existence of discriminant validity should be judged based on whether the square root of AVE is higher than the correlation coefficient value between the two-factor constructs. This study uses the maximum likelihood method to estimate the model, and  $\chi^2/df$ , RMSEA, SRMR, NFI, CFI, TLI, IFI indicators are used to verify the model fitting degree in this paper.

It can be seen from Table 6 that the reliability of A, R, C, HP, CK, ES, and MC is 0.886, 0.748, 0.843, 0.767, 0.855, 0.801, and 0.832, respectively, indicating that the questionnaire has good reliability. In addition, as the result shows that  $\chi^2/df=1.330$ , RMSEA=0.040, SRMR=0.043, NFI=0.905, CFI=0.974, TLI=0.967, IFI=0.975, it indicates that the confirmatory factor analysis model fits well. The composite reliability values of A, R, C, HP, CK, ES, and MC are 0.887, 0.752, 0.843, 0.767, 0.857, 0.808, and 0.835, which are all above 0.7, and the AVE values are 0.723, 0.503, 0.644, 0.666, 0.584, and 0.629, respectively, which are above 0.5, so it shows that the questionnaire has good convergent validity.

Table 7 shows the mean and standard deviation of A, R, C, HP, CK, ES, and MC. The AVE values of A, R, C, HP, CK, ES, and MC are 0.850, 0.709, 0.802, 0.789, 0.816, 0.764, and 0.793, respectively, which are higher than their corresponding correlation coefficients, which shows that the questionnaire has good discriminant validity.

*5.2. Regression Analysis.* Regression analysis is a statistical analysis method to determine the interdependent

relationship between two variables or multiple variables. Therefore, this paper chooses regression analysis to test the hypothesis, in which gender, age, and Internet experience (in years) are used as control variables.

Table 8 shows the predictive effect of the predictor variables on the dependent variables and the magnitude of the explanation rate. Thus, the results are shown below.

Taking self-presentation as the dependent variable, it can be seen that relatedness, having a place, and expressing self-identity have a significant positive predictive effect on self-presentation ( $\beta=0.223$ ,  $p<0.01$ ;  $\beta=0.300$ ,  $p<0.001$ ;  $\beta=0.219$ ,  $p<0.01$ , respectively). In addition, the explanatory rate of the predictor variables to self-presentation is 29.4%.

Taking content sharing as the dependent variable, it is shown that having a place and expressing self-identity have a significant positive predictive effect on content sharing ( $\beta=0.260$ ,  $p<0.001$ ;  $\beta=0.179$ ,  $p<0.05$ , respectively). The explanatory rate of the predictor variables to content sharing is 15.1%.

Taking relationship formation as the dependent variable, it can be seen that relatedness and expressing self-identity have a significant positive predictive effect on relationship formation ( $\beta=0.245$ ,  $p<0.01$ ;  $\beta=0.340$ ,  $p<0.001$ , respectively). The explanatory rate of the predictor variable to relationship formation is 25.2%.

Taking browsing others' content as the dependent variable, it is shown that autonomy and coming to know the self have a significant positive predictive effect on browsing others' content ( $\beta=0.309$ ,  $p<0.001$ ;  $\beta=0.246$ ,  $p<0.01$ , respectively). The explanatory rate of the predictor variable to browsing others' content is 16.9%.

Taking meta-voicing as the dependent variable, it can be seen that relatedness, competence, coming to know the self, and expressing self-identity have a significant positive predictive effect on meta-voicing ( $\beta=0.227$ ,  $p<0.01$ ;  $\beta=0.270$ ,  $p<0.001$ ;  $\beta=0.189$ ,  $p<0.01$ ;  $\beta=0.167$ ,  $p<0.05$ , respectively). The explanatory rate of predictor variables for meta-voicing is 39.7%.

Taking communication as the dependent variable, it can be seen that the predictive effects of the various variables of psychological needs are not significant.

Taking sourcing as the dependent variable, it can be seen that competence has a significant positive predictive effect on sourcing ( $\beta=0.366$ ,  $p<0.001$ ), and the explanatory rate of the predictor variable to sourcing is 14.9%.

Taking knowledge attributes as the dependent variable, it can be shown that competence has a significant positive predictive effect on knowledge attributes ( $\beta=0.369$ ,  $p<0.001$ ), and the explanatory rate of the predictor variable to knowledge attributes is 14.4%.

## 6. Discussion

From the data results above, the following observations could be reached. One is that each of the Zhihu affordances is likely to be related to fulfilling some psychological needs. At the same time, the salient psychological needs that drive people to use Zhihu could be identified. The other is that links between individuals' psychological

TABLE 6: Reliability and convergent validity.

Constructs	Items	Loadings	Cronbach's alpha	Composite reliability	Average variance extracted
A	A1	0.835	0.886	0.887	0.723
	A2	0.859			
	A3	0.856			
R	R1	0.738	0.748	0.752	0.503
	R2	0.698			
	R3	0.691			
C	C1	0.740	0.843	0.843	0.644
	C2	0.888			
	C3	0.771			
HP	HP1	0.801	0.767	0.767	0.623
	HP2	0.777			
CK	CK1	0.779	0.855	0.857	0.666
	CK2	0.799			
	CK3	0.868			
ES	ES1	0.724	0.801	0.808	0.584
	ES2	0.790			
	ES3	0.778			
MC	MC1	0.835	0.832	0.835	0.629
	MC2	0.780			
	MC3	0.762			

$$\chi^2/df = 1.330, RMSEA = 0.040, SRMR = 0.043, NFI = 0.905, CFI = 0.974, TLI = 0.967, IFI = 0.975$$

Note: A = autonomy; C = competence; HP = having a place; CK = coming to know the self; MC = maintaining continuity of self-identity; ES = expressing self-identity; R = relatedness.

TABLE 7: Summary statistics and discriminant validity.

	A	R	C	HP	CK	ES	MC
A	<b>0.850</b>						
R	0.284 **	<b>0.709</b>					
C	0.325 ***	0.560 ***	<b>0.802</b>				
HP	0.351 ***	0.266 **	0.311 ***	<b>0.789</b>			
CK	0.282 ***	0.495 ***	0.277 **	0.393 ***	<b>0.816</b>		
ES	0.262 **	0.559 ***	0.483 ***	0.268 **	0.459 ***	<b>0.764</b>	
MC	0.256 **	0.684 ***	0.391 ***	0.306 **	0.654 ***	0.446 ***	<b>0.793</b>
Mean	4.437	3.871	4.299	4.403	3.921	3.803	3.858
S.D.	0.726	0.734	0.697	0.700	0.791	0.822	0.842

Note: \* $p < 0.05$ , \*\* $p < 0.01$ , and \*\*\* $p < 0.001$ ; the diagonal elements represent the square root of the AVE.

needs and Zhihu affordances could be identified. Details are as follows.

The need for autonomy can motivate people to use Zhihu that has the affordance of browsing others' content. However, it has no relationship with self-presentation, content sharing, relationship formation, and sourcing. This means that although Zhihu users do have the freedom to determine what content to browse and when, there are some restrictions that limit users' freedom to show themselves, share content, and establish relationships. In the discussion area of the Zhihu platform, we can see many people complaining about the privacy problems of Zhihu. Some people think that it is strongly required that Zhihu set a function that can determine whether users are willing to disclose their dynamics, including the behavior of liking, collecting, and

following. Otherwise, when users' friends or other people who follow the users click into the accounts, they can easily know what the users are doing or thinking in recent time. Mainly when subscribing to sensitive and private topics, users feel that their privacy has been violated. In addition, although the Internet search index improves the accuracy of predicting users' preferences and behavior [78], the recommender system used in the application may also disclose users' privacy [79]. In this case, due to social norms and impression management concerns, users may feel that their behavior cannot be entirely determined by themselves and finally think that the platform cannot meet their needs for autonomy. This view is further confirmed by the comparative study between Facebook and Twitter. Facebook's real name policy makes many users feel limited in expressing

TABLE 8: Regression results.

	SP	CS	RF	BOC	MV	COM	SO	KA
	Psychological needs							
A	0.027 (0.089)	0.053 (0.097)	-0.018 (0.088)	0.309*** (0.079)			-0.008 (0.079)	
R	0.223** (0.106)		0.245** (0.097)	-0.075 (0.084)	0.227** (0.079)	0.051 (0.130)		
C					0.270*** (0.078)		0.366*** (0.084)	0.369*** (0.055)
HP	0.300*** (0.092)	0.260*** (0.101)						
CK				0.246** (0.078)	0.189** (0.066)			
ES	0.219** (0.084)	0.179* (0.088)	0.340*** (0.085)		0.167* (0.067)	0.114 (0.116)		
MC	0.013 (0.088)	0.092 (0.087)						
	Controls							
Gender	0.079 (0.123)	-0.021 (0.134)	-0.106 (0.125)	-0.095 (0.111)	0.031 (0.095)	0.050 (0.171)	-0.009 (0.113)	0.052 (0.077)
Age	0.047 (0.089)	-0.035 (0.097)	0.142* (0.090)	-0.004 (0.080)	-0.044 (0.069)	-0.085 (0.122)	-0.126 (0.080)	-0.056 (0.055)
IE	-0.102 (0.122)	-0.046 (0.132)	0.019 (0.125)	0.065 (0.110)	-0.101 (0.093)	-0.075 (0.169)	0.092 (0.108)	-0.117 (0.073)
R <sup>2</sup>	0.321	0.180	0.274	0.193	0.417	0.043	0.170	0.161
Adj R <sup>2</sup>	0.294	0.151	0.252	0.169	0.397	0.019	0.149	0.144

Note: \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , standardized coefficients (standard errors); IE=Internet experience; Gender: 1 = male, 2 = female; SP = self-presentation; CS = content sharing; RF = relationship formation; BOC = browsing others' content; MV = meta-voicing; COM = communication; SO = sourcing; KA = knowledge attributes.

themselves, while Twitter's anonymity system reduces users' social pressure and enables people to express themselves more freely [10, 80, 81].

The need for relatedness can motivate the use of Zhihu that has affordances of self-presentation, relationship formation, and meta-voicing, but there is no significant link from relatedness to browsing others' content and communication. Although browsing others' content and communication are also likely to help users develop their relationships with others, our results show that these affordances are not able to meet the psychological needs of relatedness. This may be because when users are provided with various affordances that can meet the same psychological needs, they will choose the one that can directly meet their psychological needs. In addition, compared to social platforms such as Wechat, OKC platforms such as Zhihu have weaker functions in chatting, which may also be why people are less likely to use chatting on Zhihu to meet needs for relatedness. According to results, the features which enable affordances of self-presentation, relationship formation, and meta-voicing are used to meet needs for relatedness on Zhihu.

The need for competence is significantly related to affordances of meta-voicing, sourcing, and knowledge attributes. As a successful OKC platform, the functions provided by Zhihu, such as commenting on posts, collecting and liking, voting for posts, suggesting questions, and searching answers, as well as the high-quality content of Zhihu, meet people's demands for applying and expanding their knowledge. The emergence of OKC platforms has changed the way people obtain information. Every online interaction or behavior of searching is the process of

ingesting content through the Internet. This method has brought many positive effects, such as a variety of content and quick access to information, but there will also be some negative effects, for example, how to filter meaningful information which indeed enables users to acquire competence. In order to achieve substantial development, Zhihu should avoid information redundancy and maintain high-quality content to win everyone's favor in the era of information overload.

The result supports the link between the need for having a place and the use of self-presentation and content sharing affordances. This shows that people create their own space by posting self-related content, writing the column, or sharing others' posts. Zhihu heavy users may invest a lot of time or money in engaging in Zhihu activities. In this way, people are able to be immersed in the environment they personalize for their own. As a result, the need for having a place can motivate the use of Zhihu that has affordances of self-presentation and content sharing.

The need for coming to know the self can also drive people to use Zhihu that has affordances of browsing others' content and meta-voicing. People can establish self-identity by receiving feedback from others and comparing themselves with other people [72, 73]. On the one hand, Zhihu can establish an effective feedback mechanism through such functions as liking and commenting so that users can constantly improve their answers, obtain new ideas, and may find the direction they are interested in and explore their potential. On the other hand, users will compare with their own articles or ideas by browsing other people's articles or opinions. Thus, they can find their own shortcomings and enhance their self-awareness. For example, by communicating with other users in the

comment area, users may continue to be inspired and feel the differences between themselves and others to form a clearer understanding of themselves.

The result shows that individuals who have a great need to express self-identity may use Zhihu that provides affordances of self-presentation, content sharing, relationship formation, and meta-voicing. However, the affordance of communication could not fulfill people's psychological need for expressing self-identity. By joining an online community or following others (enable the affordance of relationship formation) on Zhihu, individuals are easy to express self-identity by establishing connections with other people. However, when users are high on the need for communicating their identities, the function of chatting may not be as rich as uploading their content, sharing others' content, commenting, or voting for others' posts (enable the self-presentation, content sharing, and meta-voicing affordances). In addition, as we mentioned before, compared to Wechat in China, the chatting function of Zhihu is not its superior function. Therefore, individuals are more likely to use self-presentation, content sharing, relationship formation, and meta-voicing to express self-identity.

Finally, the need for maintaining continuity of self-identity plays a nonsignificant role in driving Zhihu use.

We expect that the realization of this psychological need may only be a by-product of using Zhihu. Individuals could express their identity by posting their own content or sharing others' content. These constitute the user's previous and current experiences and records. Unless this material is specifically deleted, the information will be persistent, which provides continuity in time and a retrospective perspective for self-identity.

## 7. Implications and Future Research

*7.1. Theoretical Implications.* The characteristics of a particular OKC application provide insights into the affordances provided by that application. In addition, from the logic of NAF, this paper provides the mapping of the psychological needs that motivate using the application to its affordances. Therefore, people can predict their psychological needs based on the features provided by the OKC application. On the contrary, based on the people's level of psychological needs, it is possible to predict which affordances of the OKC application they may use.

Taking Zhihu as an example, this paper determines the psychological needs that drive the use of Zhihu from the perspective of the NAF model. Through the features of Zhihu, Zhihu affordances are identified, so the connection between affordances and psychological needs is established. Through the data collection and analysis, the results show that the salient psychological needs that motivate users to use Zhihu are autonomy, relatedness, competence, having a place, coming to know the self, and expressing self-identity, and these psychological needs are fulfilled by the affordances of browsing others' content (for autonomy), self-presentation, relationship formation and meta-voicing (for relatedness), meta-voicing, sourcing and knowledge attributes

(for competence), self-presentation and content sharing (for having a place), browsing others' content and meta-voicing (for coming to know the self), and self-presentation, content sharing, relationship formation, and meta-voicing (for expressing self-identity). These results demonstrate how to use the NAF model to determine the relationship between a specific OKC platform and psychological needs.

*7.2. Practical Implications.* From the fact that Wukong Q&A has invested billions in the construction of the community although it still fell to the ground, different platform affordances could create different community ecology. The research of this topic can guide more platforms to design more effective OKC applications under the premise of unifying user needs and organizational characteristics, which could lead to avoiding waste of social resources and improving the success rate of Internet innovation. Outwardly, the digital platform system has a certain degree of imitability and reproducibility. In fact, it contains differences in corporate value pursuit, user community, and core capabilities. The platform system is a collection of affordances generated during the interaction between the organizational characteristics of the platform enterprise and the needs of users. The platform must clearly understand the dominance and decisive role of these organizational characteristics in operations. Different organizational characteristics naturally require different digital technologies and combination options to achieve differentiated functions, interfaces, and processes and produce different user interaction effects. In addition to the social affordances provided by the platform, content platforms such as Zhihu should also focus on the affordances of knowledge content, that is, whether the content resources meet the platform positioning and user needs. Furthermore, platform strategies should be used to guide content production behavior and user consumption behavior to build a good user-content collaborative ecosystem.

*7.3. Future Research.* NAF model can be extended to design science research. For example, our results can be used to provide guidance on how to develop effective OKC application features. These features should be able to meet the natural psychological needs of users. In addition, although we have studied how each type of affordance meets specific psychological needs, it is also possible that affordances have joint complementary effects. Therefore, future research can explore whether a specific combination of affordances can provide better means to meet the psychological needs or whether a specific combination of characteristics can provide specific affordances in a better way. Third, it is unclear whether OKC applications should strive to provide multiple functions to meet a single demand or provide multiple functions to meet multiple needs in order to succeed. Our empirical research in the context of Zhihu shows that although Zhihu provides a variety of affordances to meet the needs for relatedness, only three of them may be significantly related to the need for relatedness. Although other affordances may also help individuals develop relationships with



others, our results seem to indicate that this is not why users use them. This may be because when individuals are provided with multiple affordances that can meet the same psychological needs, they will choose the affordance that most directly meets the psychological needs. This is worthy of further study. Finally, we mentioned earlier that compared to other Q&A platforms, Zhihu has carried out in-depth content production and provided online and offline knowledge products. Whether users meet their psychological needs online or offline may depend on whether OKC provides better functions to meet their needs than offline. These questions are interesting directions for future design scientific research and provide potential ways for future OKC research to reveal how to attract users effectively.

### Data Availability

The data are available from the corresponding author upon request.

### Conflicts of Interest

The authors declare no conflicts of interest related to this article.

### References

- [1] I. Matar, "The integration of knowledge management into the Lebanese graphic design business sector," *Journal of System and Management Sciences*, vol. 8, no. 1, pp. 61–82, 2018.
- [2] Z. K. Lin, F. J. Liu, and N. Zhao, "Research review on intrinsic mechanism of group collaboration in online knowledge community," *Information Science*, vol. 37, no. 6, 2019.
- [3] J. N. Qiu, Y. Zhang, and K. Xu, "Research on the co-evolution of knowledge and social system based on individual heterogeneity in OKC," *Operations Research and Management Science*, vol. 27, no. 5, pp. 119–129, 2018.
- [4] S. X. Zeng, *The Research On Wukong Q&A Content Operation*, Hebei Technology University, Tianjin, China, 2016.
- [5] L. H. Huang and Q. H. Zhu, "An analysis of wiki entry features and user contribution behavior: the case of baidu baike," *Journal of Library Science in China*, vol. 39, no. 203, pp. 79–88, 2013.
- [6] A. Halfaker, R. S. Geiger, J. T. Morgan, and J. Riedl, "The rise and decline of an open collaboration system: how Wikipedia's reaction to popularity is causing its decline," *American Behavioral Scientist*, vol. 57, no. 5, pp. 664–688, 2013.
- [7] R. E. Rice, S. K. Evans, K. E. Pearce, and A. Sivunen, "Organizational media affordances: perationalization and associations with media use," *Journal of Communication*, vol. 67, no. 1, pp. 106–130, 2017.
- [8] A. Ghazawneh and O. A. Henfridsson, "Paradigmatic analysis of digital application market places," *Journal of Information Technology*, vol. 30, no. 3, pp. 198–208, 2015.
- [9] R. W. Ongus and C. M. Nyamboga, "Collecting development practices in using information technology: a comparative study," *Journal of Logistics, Informatics and Service Science*, vol. 6, no. 2, pp. 1–22, 2019.
- [10] E. Karahanna, S. X. Xu, Y. Xu, and N. Zhang, "The needs-affordances-features perspective for the use of social media," *MIS Quarterly*, vol. 42, no. 3, pp. 737–756, 2018.
- [11] G. G. Parker and M. W. V. Alstynne, "Two-sided network effects: a theory of information product design," *Management Science*, vol. 51, pp. 1494–1504, 2005.
- [12] Q. Zhong, X. F. Yang, and Z. Q. Wu, "A review of value co-creation in platform ecosystems," *System Engineering-Theory and Practice*, vol. 11, pp. 1–16, 2020.
- [13] S. Luo, T. Yoshita, and Y. Wang, "Online knowledge community: conceptual clarification and a CAS view for its collective intelligence," *Knowledge Science Engineering and Management*, Springer, Berlin, Germany, 2013.
- [14] Y. Lu, C. Xiang, and B. Wang, "What affects information systems development team performance? An exploratory study from the perspective of combined socio-technical theory and coordination theory," *Computers in Human Behavior*, vol. 27, no. 2, pp. 811–822, 2011.
- [15] U. Cress, I. Feinkohl, and J. Jirschtzka, "Mass collaboration as coevolution of cognitive and social systems," *Mass Collaboration and Education*, Springer International Publishing, New York, NY, USA, pp. 85–104, 2016.
- [16] M. Alavi, D. E. Leidner, and W. Q. Zheng, "Knowledge management and knowledge management system: the conceptual foundation and research topic," *MIS Quarterly*, vol. 25, pp. 107–136, 2001.
- [17] A. Beldad, M. D. Jong, and M. Steehouder, "How shall I trust the faceless and the intangible? A literature review on the antecedents of online trust," *Computers in Human Behavior*, vol. 26, no. 5, pp. 857–869, 2010.
- [18] T. Guan, L. Wang, J. Jin, and X. Song, "Knowledge contribution behavior in online Q&A communities: an empirical investigation," *Computers in Human Behavior*, vol. 81, pp. 137–147, 2018.
- [19] H. M. Lai, C. P. Chen, and Y. F. Chang, "Determinants of knowledge seeking in professional virtual communities," *Behaviour and Information Technology*, vol. 33, no. 5, pp. 522–535, 2014.
- [20] X. Zhao and S. M. Huang, "A comparative analysis between the antecedents of PVC knowledge seeking and antecedents of PVC knowledge contribution," *Journal of Intelligence*, vol. 36, no. 12, pp. 180–185, 2017, 137.
- [21] S. Faraj, S. Kudaravalli, and M. M. Wasko, "Leading collaboration in online communities," *MIS Quarterly*, vol. 39, no. 2, pp. 393–412, 2015.
- [22] J. J. Gibson, *The Ecological Approach to Visual Perception*, pp. 227–235, Lawrence Erlbaum Associates, Florence, KY, USA, 1986.
- [23] A. Majchrzak and S. Faraj, "The contradictory influence of social media affordances on online communal knowledge sharing," *Journal of Computer-Mediated Communication*, vol. 19, no. 1, pp. 38–55, 2013.
- [24] A. L. Fayard and J. Weeks, "Affordances for practice," *Information and Organization*, vol. 24, no. 4, pp. 236–249, 2014.
- [25] H. Postigo, "The socio-technical architecture of digital labor: converting play into YouTube money," *New Media & Society*, vol. 18, no. 2, pp. 332–349, 2016.
- [26] Y. Sun, S. J. He, R. A. Shang, and J. D. Fu, "Research on the mechanism of corporate social work platforms affecting employees' improvisation ability-based on the perspective of online social networks," *Management World*, vol. 35, no. 3, pp. 157–168, 2019.
- [27] J. W. Treem and P. M. Leonardi, "Social media use in organizations: exploring the affordances of visibility, editability, persistence, and association," *Communication Yearbook*, vol. 36, pp. 143–189, 2012.



- [28] M. Zhang, *Research on the Influence of Function Affordance on the Traffic Import from Social Media to E-Commerce*, Harbin Institute of Technology, Harbin, China, 2020.
- [29] J. Fox and B. McEwan, "Distinguishing technologies for social interaction: the perceived social affordances of communication channels scale," *Communication Monographs*, vol. 84, no. 3, pp. 298–318, 2017.
- [30] X. Y. Dong, *The Impact of IT Affordance and Social Ties on Social Commerce Purchase Intention*, Harbin Institute Of Technology, Harbin, China, 2018.
- [31] Z. D. Pan and Y. S. Liu, "What is 'new'? The power trap in the discourse of 'new media' and the researcher's theoretical introspection: an interview with Professor Pan Zhongdang," *News and Communications Review*, vol. 1, pp. 2–19, 2017.
- [32] Z. A. Zhang and J. L. Huang, "Research and implications on the affordance of internet platforms from the perspective of communication studies," *New Writing*, vol. 10, pp. 87–95, 2020.
- [33] E. L. Deci and R. M. Ryan, *Intrinsic Motivation and Self-Determination in Human Behavior*, Plenum Group, London, UK, 1985.
- [34] J. M. Hermans and E. Hermans-Jansen, *Self-Narratives: The Construction of Meaning in Psychotherapy*, The Guilford Press, New York, NY, USA, 1995.
- [35] H. J. Hermans, "The dream in the process of valuation: a method of interpretation," *Journal of Personality and Social Psychology*, vol. 53, no. 1, pp. 163–175, 1987.
- [36] E. L. Deci and R. M. Ryan, "The support of autonomy and the control of behavior," *Journal of Personality and Social Psychology*, vol. 53, no. 6, pp. 1024–1037, 1987.
- [37] C. Burk, "The collecting instinct," *Pedagogical Seminary*, vol. 7, no. 2, pp. 179–207, 1900.
- [38] H. Dittmar, *The Social Psychology Of Material Possessions: To Have is to Be, Hemel Hempstead*, Harvester Wheatsheaf, Hertfordshire, UK, 1992.
- [39] J. D. Porteous, "Home: the territorial core," *Geographical Review*, vol. 66, no. 4, pp. 383–390, 1976.
- [40] S. P. Shi, X. Y. He, J. Zhang, J. Hong, and Y. Wu, "Value realization of internet start-ups from the perspective of affordance: a case study based on zhong you bang," *Management Case Studies Journal*, vol. 13, no. 3, pp. 315–330, 2020.
- [41] A. Davis, J. Murphy, D. Owens, D. Khazanchi, and I. Zigurs, "Avatars, people, and virtual worlds: foundations for research in metaverses," *Journal of the Association for Information Systems*, vol. 10, no. 2, pp. 91–117, 2009.
- [42] D. Halpern and J. Gibbs, "Social media as a catalyst for online deliberation? Exploring the affordances of Facebook and YouTube for political expression," *Computers in Human Behavior*, vol. 29, no. 3, pp. 1159–1168, 2013.
- [43] I. Junglas, L. Goel, C. Abraham, and B. Ives, "The social component of information systems-how sociability contributes to technology acceptance," *Journal of the Association for Information Systems*, vol. 14, no. 10, pp. 585–616, 2013.
- [44] L. Nardon and K. Aten, "Valuing virtual worlds: the role of categorization in technology assessment," *Journal of the Association for Information Systems*, vol. 13, no. 10, pp. 772–796, 2012.
- [45] J. H. Kietzmann, K. Hermkens, I. P. McCarthy, and B. S. Silvestre, "Social media? Get serious! Understanding the functional building blocks of social media," *Business Horizons*, vol. 54, no. 3, pp. 241–251, 2011.
- [46] M. Mesgari and F. Faraj, "Technology affordances: the case of Wikipedia," in *Proceedings of the 18th Americas Conference on Information Systems*, Seattle, WA, USA, 2012.
- [47] L. Goel, N. Johnson, I. Junglas, and B. Ives, "Predicting users' return to virtual worlds: a social perspective," *Information Systems Journal*, vol. 23, no. 1, pp. 35–63, 2013.
- [48] S. Faraj, S. L. Jarvenpaa, and A. Majchrzak, "Knowledge collaboration in online communities," *Organization Science*, vol. 22, no. 5, pp. 1224–1239, 2011.
- [49] Z. M. Zhu, B. Delphine, and G. Iryna, "A multi-dimensional model for assessing the quality of answers in social Q&A sites," in *Proceedings of the 14th International Conference on Information Quality*, Potsdam, Germany, November 2009.
- [50] S. Kim and S. Oh, "Users' relevance criteria for evaluating answers in asocial Q&A site," *Journal of the American Society for Information Science and Technology*, vol. 60, no. 4, pp. 716–727, 2009.
- [51] J. Gibbs, N. Rozaidi, and J. Eisenberg, "Overcoming the 'ideology of openness': probing the affordances of social media for organizational knowledge sharing," *Journal of Computer-Mediated Communication*, vol. 19, no. 1, pp. 102–120, 2013.
- [52] S. Cai, H. R. Shi, X. Fu, and X. Chen, "Research on the factors influencing the sales of paid knowledge products: taking Zhihu live as an example," *Journal of Management in Engineering*, vol. 33, no. 3, pp. 71–83, 2019.
- [53] H. S. Cai, "Examining economic law from the perspective of interests," *Social Scientist*, vol. 7, 2008.
- [54] B. H. Zou and H. Luo, "Knowledge payment—a knowledge dissemination model centered on openness, sharing and payment," *New Media Research*, vol. 3, no. 11, pp. 110–112, 2017.
- [55] K. T. Tian, W. Bearden, and G. L. Hunter, "Consumer need for uniqueness: scale development and validation," *Journal of Consumer Research*, vol. 28, no. 1, pp. 50–67, 2001.
- [56] C. Shah and J. Pomerantz, "Evaluating and predicting answer quality in community QA," in *Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, Geneva, Switzerland, July 2010.
- [57] Y. Q. Hou and M. Xiao, "The development of socialized question-and-answer community and its communication characteristics: taking Zhihu as an example," *Northern Media Research*, vol. 4, pp. 37–41, 2017.
- [58] Zhihu.Com, 2021, <https://www.zhihu.com>.
- [59] N. N. Yin, "Research on the development of Zhihu APP knowledge service-based on SWOT-AHP analysis," *Today's Massmedia*, vol. 29, no. 3, pp. 34–36, 2021.
- [60] iResearch, *China Knowledge Marketing White Paper—Taking Zhihu as an example in 2018*, iResearch Publications, Shanghai, China, 2018.
- [61] E. L. Deci, *Why We Do What We Do: Understanding Self Motivation*, Penguin Books, New York, NY, USA, 1995.
- [62] E. L. Deci and R. M. Ryan, *Handbook of Self-Determination Research*, University of Rochester Press, Rochester, NY, USA, 2002.
- [63] E. L. Deci and R. M. Ryan, "The 'what' and 'why' of goal pursuits: human needs and the self-determination of behavior," *Psychological Inquiry*, vol. 11, no. 4, pp. 227–268, 2000.
- [64] M. A. Jenkins-Guarnieri, S. L. Wright, and L. M. Hudiburgh, "The relationships among attachment style, personality traits, interpersonal competency, and Facebook use," *Journal of Applied Developmental Psychology*, vol. 33, no. 6, pp. 294–301, 2012.
- [65] J. P. Seder and S. Oishi, "Ethnic/racial homogeneity in college students' Facebook friendship networks and subjective well-

- being,” *Journal of Research in Personality*, vol. 43, pp. 438–443, 2009.
- [66] J. J. Bauer and D. P. McAdams, “Competence, relatedness, and autonomy in life stories,” *Psychological Inquiry*, vol. 11, no. 4, pp. 276–279, 2000.
- [67] R. W. White, “Motivation reconsidered: the concept of competence,” *Psychological Review*, vol. 66, no. 5, pp. 297–333, 1959.
- [68] H. Barki, G. Paré, and C. Sicotte, “Linking IT implementation and acceptance via the construct of psychological ownership of information technology,” *Journal of Information Technology*, vol. 23, no. 4, pp. 269–280, 2008.
- [69] T. M. Harrison and B. Barthel, “Wielding new media in web 2.0: exploring the history of engagement with the collaborative construction of media products,” *New Media & Society*, vol. 11, no. 1/2, pp. 155–178, 2009.
- [70] J. L. Pierce, T. Kostova, and K. T. Dirks, “Toward a theory of psychological ownership in organizations,” *Academy of Management Review*, vol. 26, no. 2, pp. 298–310, 2001.
- [71] J. L. Pierce, T. Kostova, and K. T. Dirks, “The state of psychological ownership: integrating and extending a century of research,” *Review of General Psychology*, vol. 7, no. 1, pp. 84–107, 2003.
- [72] L. Festinger, “A theory of social comparison processes,” *Human Relations*, vol. 7, no. 2, pp. 117–140, 1954.
- [73] G. H. Mead, *Mind, Self and Society*, University of Chicago Press, Chicago, IL, USA, 1934.
- [74] E. L. Deci and R. M. Ryan, “A motivational approach to self: integration in personality,” in *Proceedings of the Nebraska Symposium on Motivation*, vol. 38, pp. 237–288, University of Nebraska Press, Lincoln, NE, USA, 1991.
- [75] R. F. DeVellis, *Scale Development Theory and Applications*, Sage, London, UK, 1991.
- [76] J. F. Hair, W. C. Black, B. J. Babin, and R. E. Anderson, *Multivariate Data Analysis: A Global Perspective*, Prentice-Hall, Upper Saddle River, NJ, USA, 2010.
- [77] C. Fornell and D. F. Larcker, “Evaluating structural equation models with unobservable variables and measurement error,” *Journal of Marketing Research*, vol. 18, no. 1, pp. 39–50, 1981.
- [78] Y. Li, M. Xu, X. Wen, and D. Guo, “The role of internet search index for tourist volume prediction based on GDFM model,” *Tehnicki vjesnik-Technical Gazette*, vol. 27, no. 2, pp. 576–582, 2020.
- [79] D. A. Sitar-Tăut, D. Mican, and C. Mare, “Customer behavior in the prior purchase stage—information search versus recommender systems,” *Economic Computation and Economic Cybernetics Studies and Research*, vol. 54, no. 3, pp. 59–76, 2020.
- [80] B. A. Huberman, D. M. Romero, and F. Wu, *Social Networks that Matter: Twitter under the Microscope*, Cornell University, Ithaca, NY, USA, 2008, <http://ssrn.com/abstract=1313405> or <http://dx.doi.org/10.2139/ssrn.1313405>.
- [81] D. J. Hughes, M. Rowe, M. Batey, and A. Lee, “A tale of two sites: twitter vs. Facebook and the personality predictors of social media usage,” *Computers in Human Behavior*, vol. 28, no. 2, pp. 561–569, 2012.

## Retraction

# Retracted: Cluster-Based Mutual Fund Classification and Price Prediction Using Machine Learning for Robo-Advisors

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] X. Chen, S. Ye, and C. Huang, "Cluster-Based Mutual Fund Classification and Price Prediction Using Machine Learning for Robo-Advisors," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 4984265, 14 pages, 2021.

## Research Article

# Cluster-Based Mutual Fund Classification and Price Prediction Using Machine Learning for Robo-Advisors

Xiaofei Chen, Shujun Ye, and Chao Huang 

Beijing Jiaotong University, School of Economics and Management, Beijing, China

Correspondence should be addressed to Chao Huang; [hchao@bjtu.edu.cn](mailto:hchao@bjtu.edu.cn)

Received 15 October 2021; Revised 6 November 2021; Accepted 18 November 2021; Published 17 December 2021

Academic Editor: José Alfredo Hernández-Pérez

Copyright © 2021 Xiaofei Chen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The rise of FinTech has been meteoric in China. Investing in mutual funds through robo-advisor has become a new innovation in the wealth management industry. In recent years, machine learning, especially deep learning, has been widely used in the financial industry to solve financial problems. This paper aims to improve the accuracy and timeliness of fund classification through the use of machine learning algorithms, that is, Gaussian hybrid clustering algorithm. At the same time, a deep learning-based prediction model is implemented to predict the price movement of fund classes based on the classification results. Fund classification carried out using 3,625 Chinese mutual funds shows both accurate and efficient results. The cluster-based spatiotemporal ensemble deep learning module shows better prediction accuracy than baseline models with only access to limited data samples. The main contribution of this paper is to provide a new approach to fund classification and price movement prediction to support the decision-making of the next generation robo-advisor assisted by artificial intelligence.

## 1. Introduction

In recent years, machine learning and deep learning have been widely used in finance to meet financial needs [1–5]. As an application of these novel techniques, robo-advisors are favored by a growing number of fund companies, and thus, the advisors have played an important role in global investments and asset allocations. Since the reform and opening up initiated 40 years ago, China has shown rapid economic growth and has become the second-largest wealth management market in the world. To satisfy residents' wealth management needs, public funds have gradually become one of the main tools for wealth management, because of their rich investment targets, professional operations, and open and transparent information environment. At the same time, with a timely supplement of technology to the financial market, the participating cost of residents to realize customized financial services has been gradually reduced. In addition, as robo-advisors are affordable to common investors, these advisors have become one of the main tools for financial institutions to carry out wealth management innovation for several advantages, e.g.,

they can manage long-term asset allocation, applying modern portfolio theory with technologies, such as big data, and implementing cloud computing. They can also automatically provide clients with fund investment suggestions through the Internet, with the consideration of investors' risk preferences, property status, and financial objectives. Therefore, the application of artificial intelligence technology outperforms traditional investment advisors by optimally dealing with several practical problems encountered by traditional advisors, e.g., cost cutting and customer reach.

There are two steps that the robo-advisors need to go through in the process of generating investment plans: fund selection and asset allocation. Fund selection can be further divided into two aspects: fund classification and fund return prediction. In China, current fund classification is generally based on the primary classification (Step 1) issued by the China Securities Regulatory Commission (CSRC). There are two main styles of funds categorized by CSRC: stock funds and bond funds. For Step 2 classification, stock funds can further be classified given their market capitalization and the style of their holdings, while bond funds can be classified based on the share of equities held in their position.



In terms of fund classification by machine learning, there are two common methods: partition clustering (represented by K-means clustering) and network clustering (represented by SOM). The partition clustering method has the merits of simpler principle, fewer parameters input, and faster convergence speed; however, the shape of circular clustering may be too simple, which would compromise the accuracy of classification results. The network clustering method can effectively deal with multidimensional clustering problems, but it is subject to dimensional disasters and the network model is relatively sensitive to the selection of parameters. In terms of fund movement prediction, ARIMA, artificial neural network (ANN), and backpropagation neural network (BP) are frequently used, while their weakness are that parameters are difficult to estimate and models are subject to overfitting.

In order to solve the above problems, we employ the Gaussian mixture clustering method (GMM) in this paper. GMM can effectively solve the problem that the cluster shape is too simple when using a simple parameter. For empirical analysis, we use data from the Chinese market. We continue our analysis using two-step classification to further distinguish the styles and characteristics of stock funds and bond funds of our collected data. Based on the GMM results, we use the spatiotemporal ensemble deep learning model to predict the short-term price movement of each category of funds. We then make full use of the idea of clustering and ensemble learning with the aim to effectively improve the implication and prediction ability of the model, especially when the access to big data is rather limited.

To compare the performance of predicting fund net asset values of our model, we compare it with several basic models, that is, residual network (ResNet, hereafter) model, long- and short-term memory network (LSTM, hereafter) model, and one-dimensional convolutional neural network (CNN, hereafter) model. We examine their performance in predicting the short-term returns of the four main classified categories in our results by employing the mean absolute error (MAE) and correlation coefficient  $R^2$  as evaluation indicators.

Our main findings are as follows. Our two-step GMM method can generate the probabilities that the funds belong to a certain category, according to their risks and returns, and thus outperform the traditional K-means model in classifying funds. Our model also improves the prediction ability, with a reduced prediction error, of fund price movements when compared with other models, i.e., ResNet, LSTM, and CNN models.

The main contributions of this paper are as follows: (1) we propose a new two-step GMM model to effectively distinguish the mutual funds in China using simple fund characteristics and (2) we construct an ensemble deep learning model to predict the short-term price movement of different categories of funds.

## 2. Literature Review

*2.1. Literature on Fund Classification.* Fund classification is the basis of fund evaluation. Different types of funds need

different analysis methods and evaluation dimensions due to their distinct characteristics such as risk and return. Thus, fund classification ensures the effectiveness and comparability of fund evaluation. There are two methods of fund classification: ex ante classification and ex post classification. The ex ante classification method determines the fund category according to its investment objectives and strategies, which are specified in the fund issuance announcement. However, in the actual operation, the specified information frequently deviates from the original agreements. Dibartolomeo et al. [6] find that after regressing the net value of the fund using the William Sharp's attribution method, more than 40% of the stock funds have misclassification, and they argue that the main reasons for the misclassification are the imprecision of the ex ante classification method and the ex post deviating manipulation by fund managers due to peer pressure. Luo et al. [7] classify 54 funds listed in China through factor and cluster analyses, and they find that nearly 40% of the funds are inconsistent with the investment style described in their prospectus. In contrast, ex post classification method specifies fund types according to their performance after fund operation and their characteristics specified in the issuance announcement. As an improvement to this classification method, Brown et al. [8] use a factor model to capture the nonlinear characteristics of fund returns and map them into the mainstream of investment managers' style to classify funds. Kim et al. [9] choose more market characteristics, through principal component analysis (PCA), and classify funds based on these newly identified variables. However, the ex post classification method also has limitations, for example, the collinearity of factors in multiple regressions.

Fortunately, the introduction of machine learning mitigates the limitations of the traditional fund classification methodology because of its ability to capture nonlinear features and its independence of data characteristics, for example, sample size, under unsupervised learning. Marathon et al. [10] classify funds by K-means clustering to find that 43% of the fund samples were inconsistent with the investment types the funds were originally described. They also find that many fund categories under traditional classification methods show very similar risk and return characteristics, thus suggesting that the introduction of classification analysis reduces the complexity of fund management. Lajbcygier et al. [11] maintain that the boundaries of funds with different styles should be continuous rather than strictly being divided. Thus, they use a flexible clustering method of fuzzy C-means and find that this method can obtain better classification results. Menardi et al. [12] employ a two-step clustering method, and the first step of which is to reduce the dimensionality of 24 fund characteristics using PCA, and the second step is to classify 1436 public funds into those 24 categories of characteristics using hierarchical clustering. For the extraction of nonlinear characteristics, Moreno et al. [13] classify 1,592 funds from the Spanish market by using self-organizing mapping neural network (SOM) and find that compared with K-means clustering, SOM can effectively reduce misclassification.

*2.2. Literature on Fund Return Prediction.* Yaser et al. (1996) argue that financial time series itself is noisy and is a nonstationary process, which means the historical information is not enough to explain the relationship between past and future returns. They also argue that the financial market is not completely unpredictable due to the existence of the price trend effect. Cao et al. (2001) maintain that financial time series can be predicted by both univariate and multivariate analyses. They argue that the input of the univariate analysis model is the time series itself, which is predicted by the autoregressive integral moving average (ARIMA) model. However, ARIMA's performance is not satisfactory, mainly because (1) it requires the parameters to be estimated *ex ante* and (2) it assumes that the time series is stable and linear, which violates reality. Concerning multivariate analysis, artificial neural network (ANN) is one of the main prediction methods due to its ability in accommodating more available information and its outstanding performance in handling nonstationary processes. Schneburg [14] uses several methods, such as MADALINE and BP models, to predict stock returns of three listed companies in the German market and reveals that the accuracy of prediction reaches 90%. Schneburg [14] also finds that the BP model performs better than other methods considered, which confirms the effectiveness of neural networks in predicting fund returns. Kimoto [15] uses a modular neural network to develop a trading strategy for Tokyo Stock Exchange and find that the modular neural network can connect more basic neural networks and help generate a more accurate prediction. For improvement, Vapnik et al. (1996) propose a new method based on the support vector machine (SVM) to improve the generalization of neural networks by solving nonlinear regression problems. In addition, Cao et al. (2001) argue that artificial neural network has overfitting problems, and extra care is needed for parameter estimations and training to get satisfactory results. Khashei et al. [16] maintain that to produce accurate results a large amount of historical data is needed, while the financial market is full of uncertainty and changes rapidly. Accordingly, a new hybrid method, which combines the advantages of artificial neural networks and fuzzy regression, is proposed to overcome the limitations of traditional artificial neural networks. Their empirical results then show that this hybrid model is an effective way to improve prediction accuracy. Liu et al. [17] and Li et al. [18] predict stock returns by employing a CNN-LSTM model and find that introducing an attention mechanism could help improve the accuracy of the CNN-LSTM model.

### 3. Data

The data used in this paper are mainly from the WIND database. The collected data are the full samples of stock funds, hybrid funds, and bond funds under the classification caliber of CSRC. QDII funds, closed-end funds, alternative funds, and monetary funds are not included in the analysis due to their peculiarity to our chosen funds. The date of April 30, 2021, is taken as the cut-off date. The funds that have been established more than two years before the cut-off

date, with a total asset value of above 50 million RMB, and have not undergone type conversion during their operation period are selected. This selection process leads to a total number of 3,625 funds.

The fund fee normally includes the daily subscription fee, daily redemption fee, management fee, and custody fee. The fund fee used in this paper is the sum of the management fee and custody fee. The value of the fee is adopted as of April 30, 2021. For fund returns, we focus on their rolling rate of return, which is defined as the ratio of the funds' compound net value on that day to the funds' compound net value  $N$  days ago, where  $N=1, 5, 20, 60, 120,$  and  $250$  days as rolling windows. The number of fundholders, shares per capita, shares that are held by institutions and individual holders, shares purchased and redeemed by the fund managers, and ratios of the market value of stocks and bonds to funds' total value are collected from their quarterly reports. Since the publication time of the quarterly reports varies across funds, we take the last trading day in January, April, July, and October of each year as the observation date to obtain the latest published data before that date. If no new data is published in the latest quarter, the value of the previous quarter will be used. Table 1 summarizes the data used in this paper.

## 4. The Model

In this section, we develop the model framework, including the clustering model and the prediction model, as shown in Figure 1. First, we cluster the fund via GMM using fund features described in the previous section, from which the dimension of the fund features is reduced using the PCA method. Then, we develop a deep learning-based prediction model to predict the fund trend based on the clustering results. Therefore, we will briefly introduce the PCA and GMM at first. The deep learning-based model is then presented in detail.

*4.1. Clustering Models.* This paper uses two models for analysis, that is, principal component analysis (PCA) and the Gaussian mixture model (GMM). In this paper, we employ PCA to reduce the dimension of the rolling yield data of funds with different maturities used in the first step of classification.

*4.1.1. PCA.* Principal components analysis (PCA) is a dimension reduction method, namely, transfers the original feature space into a brand-new feature space. It is generally used for data preprocessing. For example, assume that there are  $n$  fund samples, with each fund sample contains 2,000 features. Among these features, there exists the vast amount of noises or useless information. Therefore, PCA can be conducted to reduce noises and save computational resources. Table 2 shows the number of eigenvalues used and the reducibility using PCA. One can notice that the PCA method has a high degree of reducibility, and the higher the reducibility, the more rolling days are selected, keeping the



TABLE 1: Summary statistics.

Item	Observations	Frequency	Step of clustering
Rate of return in the past 1 day	$237 \times 3,604$	Daily	First
Rate of return in the past 5 days	$237 \times 3,604$	Daily	First
Rate of return in the past 20 days	$237 \times 3,604$	Daily	First
Rate of return in the past 60 days	$237 \times 3,604$	Daily	First
Rate of return in the past 120 days	$237 \times 3,604$	Daily	First
Rate of return in the past 250 days	$237 \times 3,604$	Daily	First
Rate of fund fee	$36 \times 4 \times 1$	As of April 30, 2021	Second
Number of fund holders	$10 \times 3,604$	Quarterly	Second
Shares per capita	$10 \times 3,604$	Quarterly	Second
Shares held by institutions (%)	$10 \times 3,604$	Quarterly	Second
Shares held by individuals (%)	$10 \times 3,604$	Quarterly	Second
Shares purchased by fund managers	$10 \times 3,604$	Quarterly	Second
Shares redeemed by fund managers	$10 \times 3,604$	Quarterly	Second
Ratios of stocks to total assets (%)	$10 \times 3,604$	Quarterly	Second
Ratios of bonds to total assets (%)	$10 \times 3,604$	Quarterly	Second

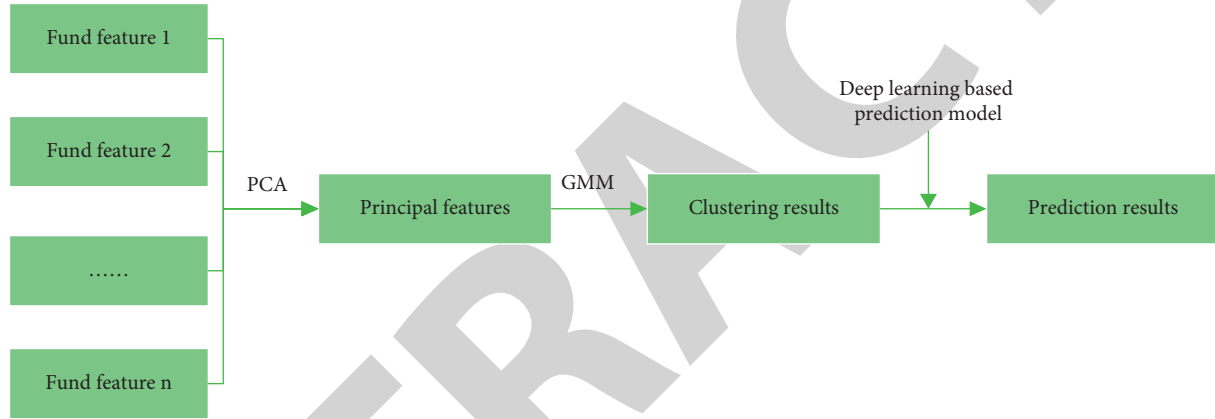


FIGURE 1: Model framework.

TABLE 2: PCA result for data used in classification.

Data	Number of eigenvalues	Reducibility (%)
Rate of return in the past 1 day	100	94.89
Rate of return in the past 5 days	100	99.19
Rate of return in the past 20 days	100	99.81
Rate of return in the past 60 days	100	99.94
Rate of return in the past 120 days	100	99.97
Rate of return in the past 250 days	100	99.98

number of eigenvalues constant. A detailed description of PCA is shown in Appendix A.

**4.1.2. GMM.** GMM clustering is the main classification model, which can give the probability that the sample belongs to a certain category. The so-called Gaussian mixture model is the combination of multiple Gaussian distributions, and it uses the likelihood function as the objective function for parameter estimation. Although a larger sample size could help improve the accuracy of the model, the increased sample size will add the complexity of the model and could thus cause overfitting problem. Therefore, we rely on Akaike information criterion (AIC) and Bayesian information

criterion (BIC) to assess the quality of classification. A detailed description of GMM is presented in Appendix B.

**4.2. Deep Learning Model.** Applying the idea of ensemble learning to the field of financial studies, we build a spatiotemporal ensemble deep learning model for fund price prediction. The rationale of this model is summarized in three steps. Given the layout and the article length, the descriptions of the components of this model are presented in Appendix C. First, encode the data that are used (i.e., historical value of the fund) to obtain the semantic information of the data. Second, insert the encoded data into three basic models, that is, residual network (ResNet), long-

and short-term memory network (LSTM), and one-dimensional convolutional neural network (CNN), to train the model and obtain the output from these three models. The detailed mechanism of ResNet, LSTM, and CNN is shown in Appendixes C.1, C.2, and C.3, respectively. Third, calculate final output; model weights are optimized by the attention mechanism, the rationale of which is shown in Appendix C.4. Those three models are employed to extract complex nonlinear correlations (ResNet), capture time correlations (LSTM), and calculate spatial correlations (CNN). In this paper, we compare the performance of our model with these three benchmark models, and the results are shown in Section 5. The detailed spatiotemporal ensemble deep learning model and the inner structure are shown in Figure 2.

## 5. Results

**5.1. Clustering Analysis.** The rolling returns of the fund in the past 1, 5, 20, 60, 120, and 250 days are used as the input of the first step GMM, in which we rely on AIC to determine the optimal number of categories. As shown in Figure 3, when the number of categories (clusters) is greater than 10, the AIC value begins to rise, which means the optimal number of classifications is 10.

We then measure our classification results of the first step GMM from multiple dimensions; the results are shown in Table 3. We start by examining the average ratio of the market value of funds' holdings of stocks and bonds to funds' total value (disclosed in their 2020 annual report) for each category. We define that for each fund when its market value of stocks accounts for less than 60% of its total value, this fund does not belong to stock funds. Thus, one can see that categories 1 and 10 are not classified as stock categories, while others are classified as stock categories. In addition, the average stock holding of category 10 is higher than that of category 1. Therefore, we define category 1 as pure bond funds while category 10 as bond-like funds. Notice that funds in category 1 have an average ratio of bonds to fund value that is greater than 100%; this is a special characteristic of bond funds that utilize leverage through repurchase agreement (repo); a similar phenomenon also occurs in Table 4. Due to a large number of funds within these two categories (more than 500 funds), we will conduct a secondary classification for these two categories, the results of which will be shown in Table 4.

Regarding the other eight categories classified as stock funds, we further define their styles according to the following criterion: a fund is defined as an industry-themed fund if more than 30% of its investments are in a certain industry or if above 50% of its holdings go to less than three specific industries. If a category of funds contains at least 50% of funds within the same industry or above 50% shares of this category is industry-themed funds, this category is defined as an industry category. Take category 4 as an example, the largest industry invested by this category is pharmaceuticals, and 91% of funds (96 out of 106) within this category invest more than 30% of their shares in the pharmaceuticals industry. Therefore, this category is defined

as the pharmaceuticals-themed (industry) category. The classification results of the categories are shown in Table 5.

For other categories, that is, categories 2, 6, 8, and 9, belonging to the category of stock but found not falling into the style of industry-themed ones, we conduct a separate classification. For category 2, its fund holdings are rather dispersed, that is, there is no representing industry. Therefore, this category is classified as a market category. Category 6 contains a large portion of funds investing in small and medium enterprises (SME), and their holdings are also relatively dispersed across industries. Therefore, category 6 is classified as an SME market category. Most funds within category 8 invest in large blue-chip industries such as banking, nonbanking finance, and food and beverage. Therefore, category 8 is classified as a large-cap style. Category 9 is found to have similar characteristics to those of category 6, while the top three industries invested by category 9's funds are pharmaceuticals, electronics, and science and technology, which are growing industries in the Chinese market. Therefore, category 9 is classified as SME growth style. The classification results of these categories are summarized in Table 6.

We continue our analysis by conducting the second step of classification for categories 1 and 10, the results of which are shown in Figure 4. According to AIC, the optimal number of clusters for each category exceeds 10, which is not practical in real-world scenarios. Bond funds are hard to be distinguished by pure quantitative measures due to their low volatility, and thus, we choose to make the minor adjustment by combining classification results with our practical experience. We can see that AIC decreases significantly when cluster number reaches 3, and bond funds are usually separated by the involvement of equity assets. As a result, we set the number of subcategories of these two categories as three to distinguish among pure bond funds, bond funded with equity, and special-purpose bond funds.

Subcategory 1 of category 1 has no excessive leverage ratio, and its share in stocks is less than 10%, which implies that subcategory 1 belongs to the primary bond category. Subcategory 2 has no positions in stocks, and its ratio of the market value of the bond to its total value is higher than that of subcategory 1. Thus, subcategory 2 can be classified as a pure bond. Subcategory 3 is classified as an institutional-customized bond category as its market value of term bonds accounts for a high proportion of its total value and its average number of holders is in single digit, implying that these funds are mainly held by several institutional investors.

Regarding category 10, the holdings of stocks of subcategory 1 accounts for more than 20%, and there is no excessive leverage. Accordingly, subcategory 1 is classified secondary bond. Subcategory 2 is classified as low dividend stock as most of its funds invest in low dividend stocks. Subcategory 3 has a very similar feature to subcategory 1 in terms of shares of stock and bond holdings. However, further investigation shows that the average market value of the funds within subcategory 3 is less than 1 billion RMB (while the figure for subcategory 1 is 1.9 billion RMB), and there is a relatively high proportion of hybrid funds within this subcategory. This indicates that subcategory 3 is mainly

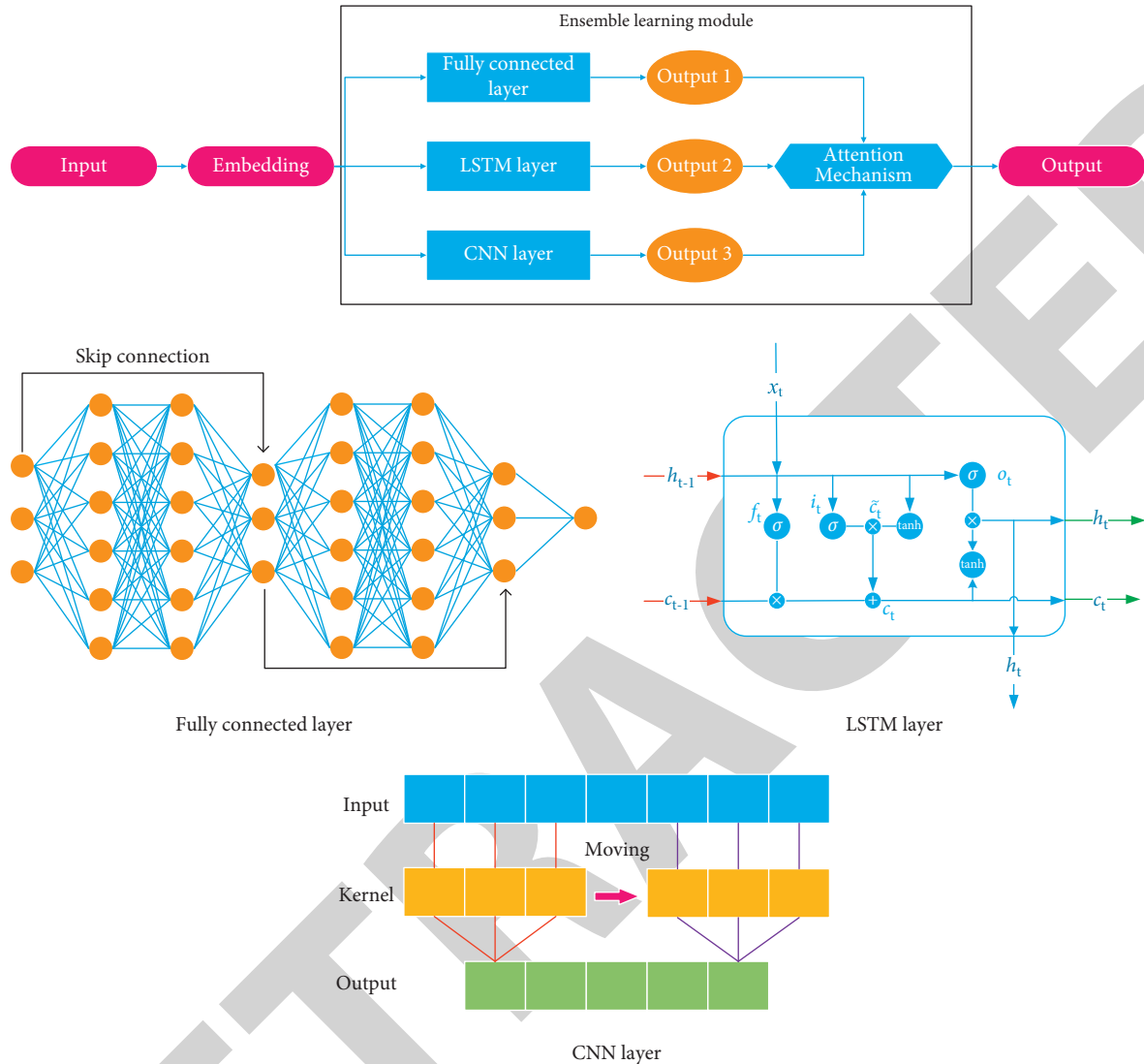


FIGURE 2: Diagram of spatial-temporal ensemble learning.

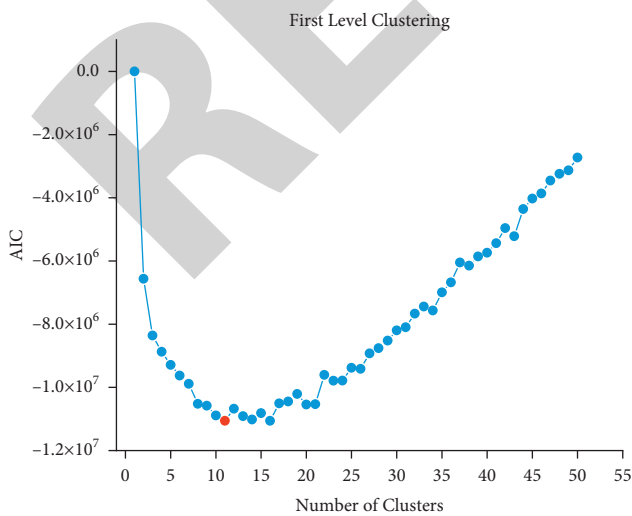


FIGURE 3: Number of main classes.

made up of fixed income and new funds established for the issuance of new stocks in the primary market. Thereby, subcategory 3 is classified as a new stock category. The secondary classification results for categories 1 and 10 are shown in Table 6.

To test the creditability of our classification results, we average the daily rate of return of the funds in each category and employ several industry style indexes (in which we select CITIC class I industry classification index as the industry index and we employ Juchao index as style index) as reference. We conduct our test by calculating the correlation coefficient between the calculated return in each category and the reference index. The results are shown in Table 7. One can see from Table 7 that category 3 has the highest correlation with CITIC food and beverage index, reaching 94%. The correlation between the return in category 4 and the CITIC pharmaceuticals index is high at 97%. The correlation between the rate of return of category 5 and the

TABLE 3: Funds category for the first step of GMM.

Category	Number of funds	Ratio of stocks to fund value (%)	Ratio of bonds to fund value (%)	Classifications
1	951	3	108	Bond
2	402	88	3	Stock
3	181	89	2	Stock
4	106	89	1	Stock
5	93	91	1	Stock
6	233	78	8	Stock
7	125	84	2	Stock
8	454	76	8	Stock
9	482	88	3	Stock
10	577	29	68	Bond-like

TABLE 4: Classification result of subclasses in categories 1 and 10.

Category	Subcategory	Share of stocks (%)	Share of bond (%)	Number of holders	Ratio of institutional investors (%)	Style
1	1	6	105	33689	65	Primary bond
	2	0	110	4019	97	Pure bond
	3	0	121	8	100	Institution customized bond
10	1	23	80	14744	90	Secondary bond
	2	78	11	71462	38	Low dividend stock
	3	23	76	14411	71	New stock

TABLE 5: Classification of categories.

Category	Number of funds	Share of the largest industry (%)	Share of largest three industries (%)	Industry	Share of the themed industry (%)	Style
3	181	38	63	Food and beverage	74	Industry
4	106	70	80	Pharmaceuticals	91	Industry
5	93	29	57	Energy and power	69	Industry
7	125	51	70	Technology	48	Industry

TABLE 6: Classification result for other stock categories.

Category	Number of funds	Share of the largest industry (%)	Share of largest three industries (%)	Industry	Share of the themed industry (%)	Style
2	402	19	41	Pharmaceuticals	12	All markets
6	233	19	37	Pharmaceuticals	45	SME-market
8	454	17	37	Food and beverage	30	Large-cap value
9	482	23	49	Pharmaceuticals	26	SME-growth

CITIC power and new energy index is 94%. The rate of return of category 7 has a high correlation coefficient with CITIC power and new energy, computers, electronics, and military industry indexes.

Table 8 summarizes the correlation between fund returns and Juchao style index for categories 2, 6, 8, and 9, which are not classified as industry-themed categories, as in Table 6. (Thus, industry indexes, such as CITIC indexes, are inappropriate to evaluate its industry classification results, while

style index, for example, Juchao index, are more suitable for the evaluation of their style classifications.) From Table 8, we can see that our classification method has high accuracy in identifying the styles of these categories. To be specific, category 2 has a higher correlation with market style indexes defined by the Juchao style index; category 6 is closer to SME market; category 8 is similar to market value, and category 9 alike SME growth. (In recent years, fund managers pay more attention to funds that are growing while less attention is

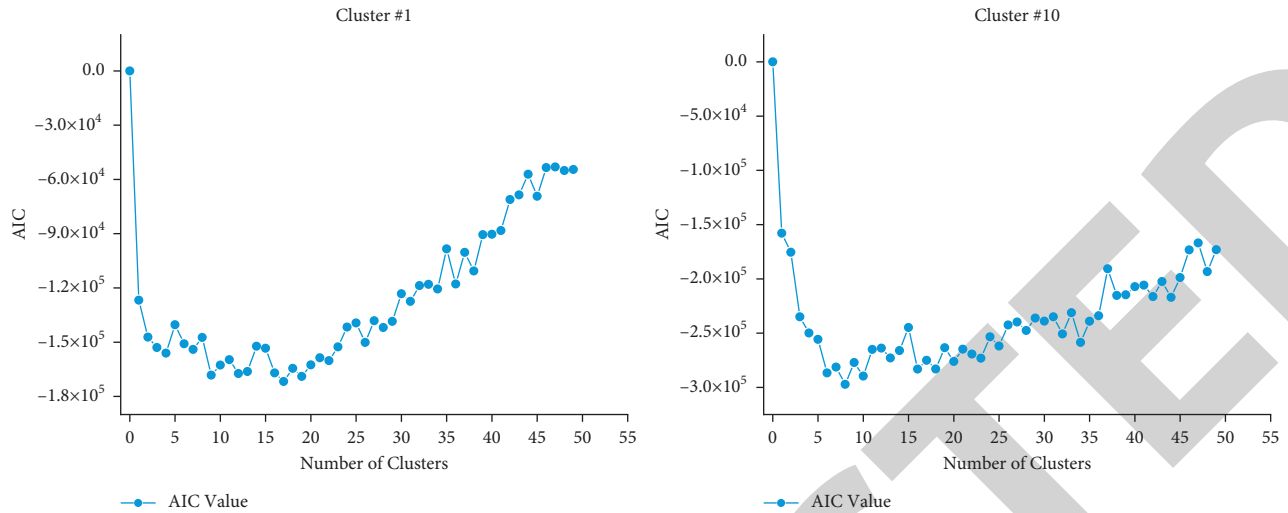


FIGURE 4: Number of subcategories of categories 1 and 10.

TABLE 7: Correlation categories and benchmark sector index (CITIC)

CITIC index	Category 3 (%)	Category 4 (%)	Category 5 (%)	Category 7 (%)
Food and beverage	<b>94</b>	74	65	58
Pharmaceuticals	82	<b>97</b>	72	67
Power and new energy	74	67	<b>94</b>	<b>85</b>
Computers	64	62	77	<b>92</b>
Communications	57	57	68	80
Electronics	64	57	81	<b>93</b>
Military	56	52	69	<b>84</b>

TABLE 8: Correlation result of other stock categories (Juchao style index).

Juchao index	Category 2 (%)	Category 6 (%)	Category 8 (%)	Category 9 (%)
Small value	82	90	88	75
Small growth	90	<b>97</b>	88	90
Medium value	83	88	<b>91</b>	75
Medium growth	<b>96</b>	<b>97</b>	<b>92</b>	<b>96</b>
Large value	73	70	<b>88</b>	61
Large growth	<b>95</b>	87	<b>92</b>	92

paid to value investing. Thus, the classification result of category 8 is less pronounced WIND database: <https://www.wind.com.cn/en/edb.html>.)

**5.2. Return Prediction Analysis.** We use the four industry-themed categories (i.e., categories 3, 4, 5, and 7) for the fund price prediction analysis. The data used is the cumulative return of all funds in these four categories from February 2019 to July 2021, totaling 604 days. The PyTorch deep learning framework is used to build the model; 80% of the data sets are used to train the model, and 20% are used to test the model. The input data is divided into 10 dimensions through embedding. The ResNet model includes 7 layers of fully connected networks, containing 16, 16, 10, 16, 16, 10, and 1 neuron. The LSTM model is made up of 1 layer and 64 neurons. The CNN model consists of three layers, and the number of filters is 8, 8, and 1, with the size of convolution kernel of 3. The batch size of all basic models is 1. The final

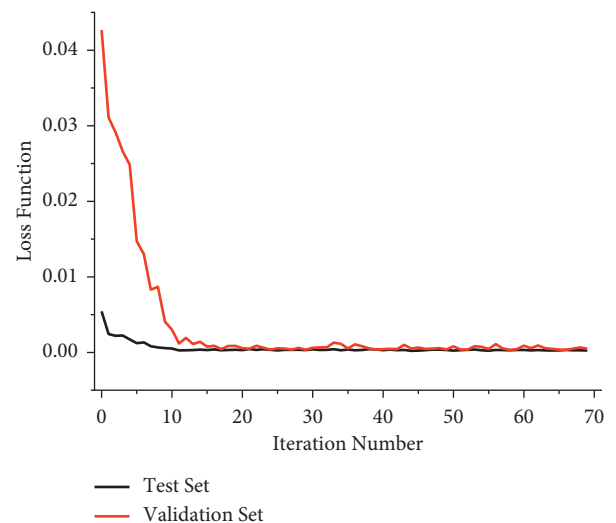


FIGURE 5: The loss function and times of iteration.

TABLE 9: Prediction result for categories 3 and 4.

	Category 3				Category 4			
	RMSE	MAE	$R^2$	WMAPE	RMSE	MAE	$R^2$	WMAPE
ResNet	0.0267	0.0213	0.9890	0.0114	0.0291	0.0220	0.9806	0.0121
LSTM	0.0331	0.0266	0.9832	0.0143	0.0333	0.0260	0.9746	0.0143
CNN	0.0308	0.0237	0.9853	0.0127	0.0389	0.0323	0.9654	0.0176
Spatiotemporal	<b>0.0219</b>	<b>0.0165</b>	<b>0.9927</b>	<b>0.0089</b>	<b>0.0262</b>	<b>0.0202</b>	<b>0.9843</b>	<b>0.0110</b>

TABLE 10: Prediction result for categories 5 and 6.

	Category 5				Category 7			
	RMSE	MAE	$R^2$	WMAPE	RMSE	MAE	$R^2$	WMAPE
ResNet	0.0329	0.0268	0.9878	0.0162	0.0466	0.0369	0.8732	0.0254
LSTM	0.0432	0.0349	0.9789	0.0211	0.0495	0.0415	0.8563	0.0286
CNN	0.0315	0.0257	0.9888	0.0155	0.0472	0.0381	0.8697	0.0263
Spatiotemporal	<b>0.0303</b>	<b>0.0246</b>	<b>0.9897</b>	<b>0.0149</b>	<b>0.0272</b>	<b>0.0211</b>	<b>0.9565</b>	<b>0.0145</b>

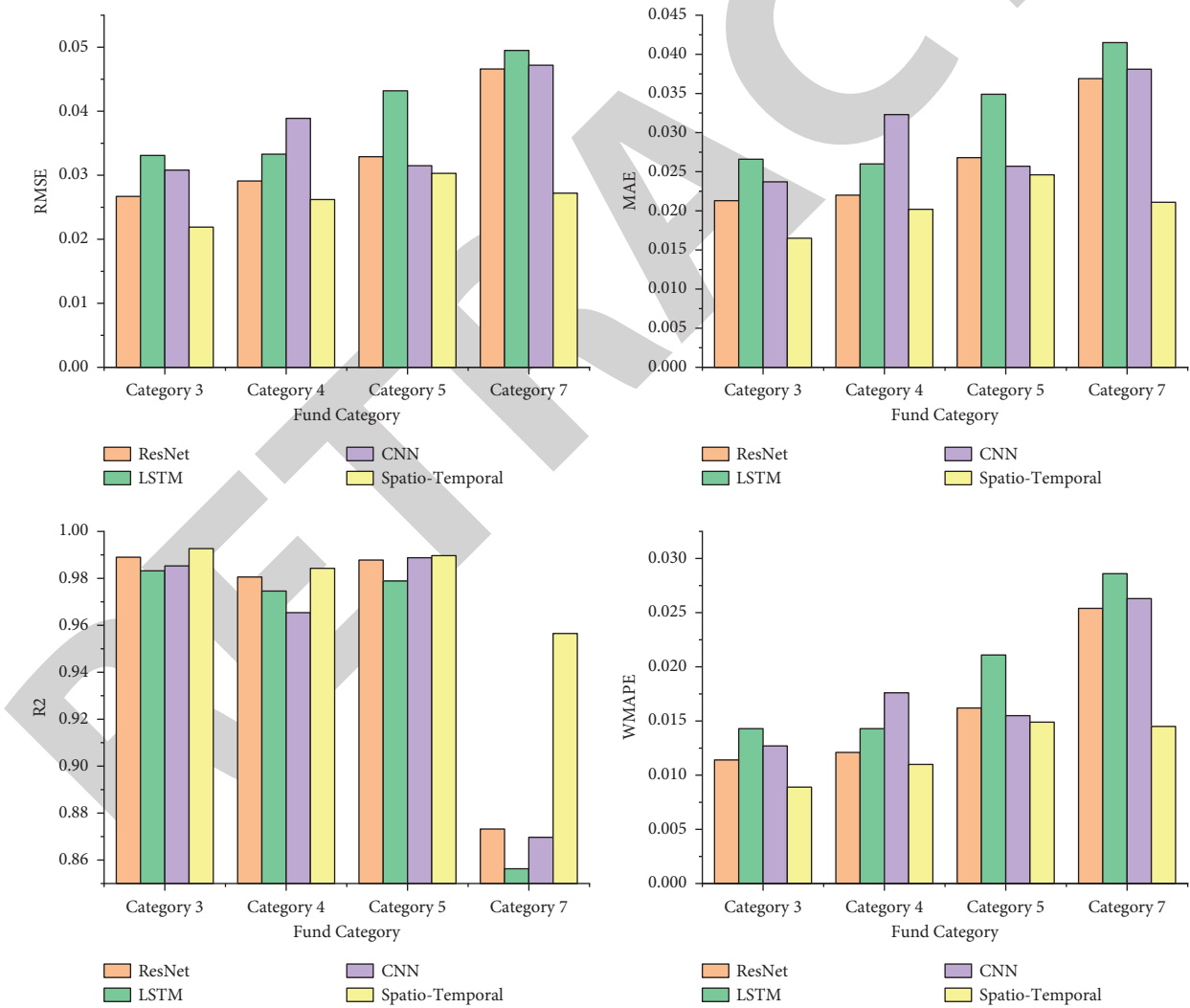


FIGURE 6: Prediction results for different categories.

output is obtained using the weighted summation of outputs of these three benchmark models (i.e., ResNet, LSTM, and CNN). The loss function curve of the model is shown in

Figure 5, from which one can notice that the loss function training and verification set become stable after the 20<sup>th</sup> iteration, and the convergence is good.



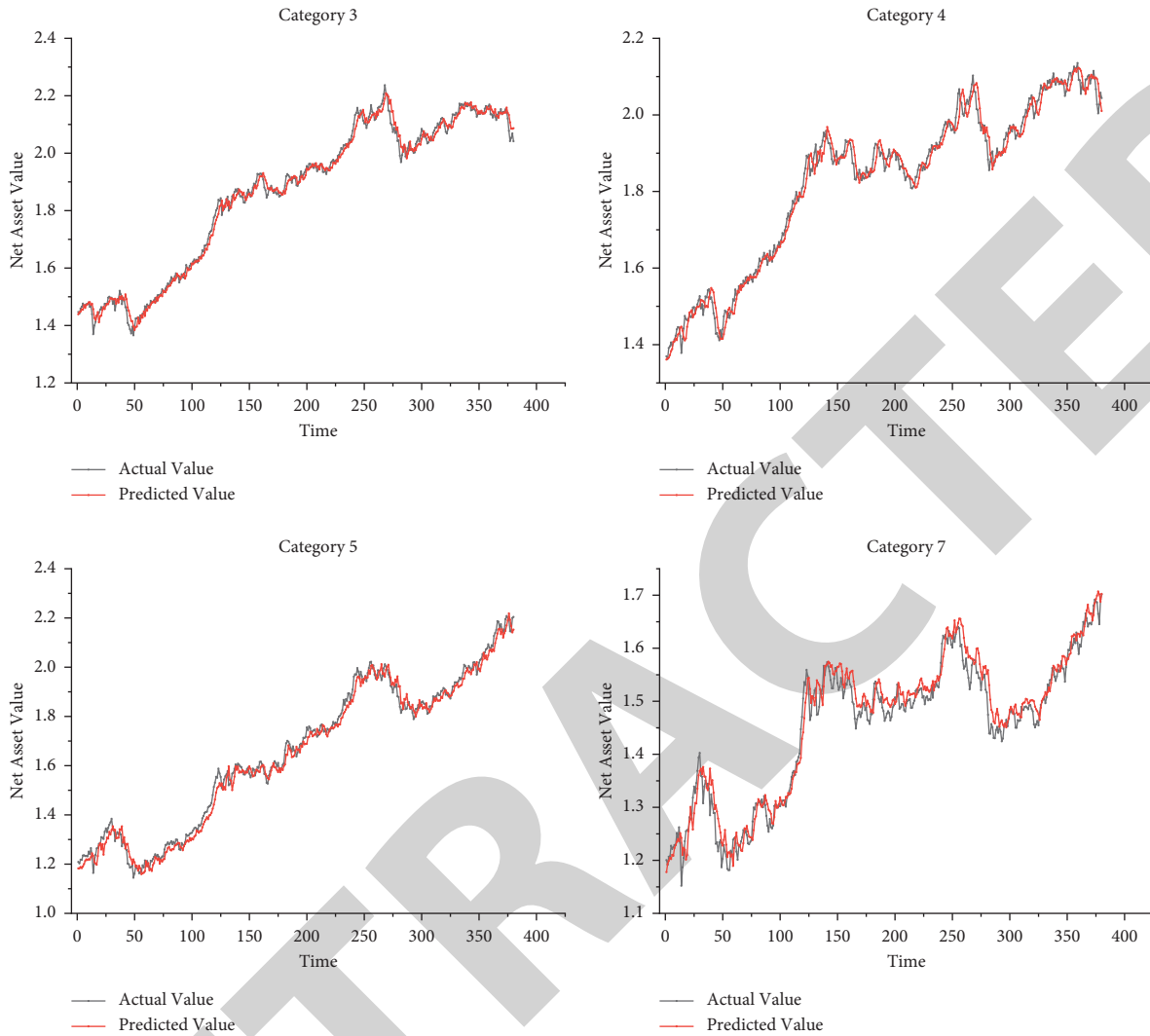


FIGURE 7: Comparison between the actual values and predicted values.

We then evaluate the performance of our model and three basic models using four indicators: RMSE, MAE,  $R^2$ , and WMAPE. The results are shown in Tables 9 and 10 and Figure 6. One can see the prediction ability of ResNet is relatively the best in all categories in question (compared with other benchmark models) based on these indicators. For example, the average value of RMSE is 0.034 for all categories, which is lowest when compared with other benchmark models, and  $R^2$  is 0.96, which is the highest among other benchmark models. Although LSTM is one of the most effective models dealing with time series problems, it does not perform well as expected, with the average value of RMSE around 0.040 and  $R^2$  around 0.95, owing to limited features available and an insufficient number of samples. The poor performance of CNN, with an average value of RMSE around 0.037 and  $R^2$  around 0.95, is due partially to its inability to record historical information. As expected, one can see that our spatiotemporal ensemble deep learning model has a better performance in predicting the fund price, which reduces the mean of RMSE to around 0.026 and improves  $R^2$  to nearly 0.98. We can thus conclude that our

model greatly improves the predictability, that is, the prediction accuracy, of fund price movement compared with these benchmark models as our model obtains the lowest (mean) values of RMSE, MAE, and WMAPE and the highest (mean) value of  $R^2$ . The improvements lie in the merits of ensemble learning, which complements the advantages of these three benchmark models.

Finally, we show the comparison between the actual values of those funds with our predicted values. The results are shown in Figure 7. We can see that the predicted values and actual values show a high degree of similarity. This means our spatiotemporal ensemble deep learning model generates a good prediction for fund price, which suggests our model can provide a proper application for robo-advisors in terms of predicting fund price movements.

## 6. Conclusion

This paper presents a novel fund price prediction tool, i.e., a spatiotemporal ensemble deep learning model, relying on fund classification, to predict the price of our selected funds,

in the Chinese market. In this paper, we propose a two-step GMM to classify the mutual funds into different categories to ensure the funds classified in the same category have similar risk and return characteristics. We then employ our proposed model to predict the short-term price movement of each fund category.

The main conclusions of this paper are summarized as follows. (1) Compared with the traditional K-means clustering method and network clustering method, our two-step GMM method can generate the probabilities that the funds belong to a certain category, (2) This paper adopts the idea of ensemble learning to improve the prediction ability of fund price movements of other models, i.e., ResNet, LSTM, and CNN models. (3) We classify funds based on their risk and return, which can effectively mitigate the problems of large fluctuations and disorders in the prediction process and thus improve the generality and application of our model.

## Appendix.

### A. PCA

Principal components analysis (PCA) is a kind of dimension reduction method and transfers the original feature space into a brand-new feature space. It is generally used for data pre-processing. For example, assume that there are  $n$  fund samples, with each fund sample with 2,000 features. Among these features, there exists the large amount of noises or useless information. Therefore, the PCA can be conducted for dimension reduction, thus reducing noises and computational resources.

Assume the feature matrix of fund samples as  $(n, m)$ , where  $n$  is the fund sample number and  $m$  is the fund feature number. The PCA can be conducted as follows.

*Step 1.* Compute the mean of each column and subtract the mean using each column to ensure the mean of each column is zero. The dimension of the feature matrix after processing is  $n \times m$ .

*Step 2.* Compute the covariance matrix of the feature matrix. The dimension of the covariance matrix is  $m \times m$ .

*Step 3.* Compute the eigenvalues and eigenvectors of the covariance matrix. The eigenvalues correspond to the eigenvectors one to one. Order the eigenvalues from largest to smallest and order the eigenvectors according to columns. Assume there are  $e$  eigenvalues, the dimension of the eigenvector matrix is  $m \times e$ .

*Step 4.* The final data can be obtained by the feature matrix after processing multiplying the eigenvector matrix. The dimension of the final data is  $n \times e$ .

*Step 5.* Choose an appropriate principal component from all principal components. Assume the  $e$  eigenvalues are  $\alpha_1, \alpha_2, \alpha_3 \dots \alpha_e$  from large to small. Then, after retaining the principal components corresponding to the first  $k$  eigenvalues, the retained variance percentage  $p$  can be obtained by the following equation:

$$p = \frac{\sum_{j=1}^k \alpha_j}{\sum_{j=1}^e \alpha_j}. \quad (\text{A.1})$$

### B. GMM

Gaussian mixture model (GMM) is a kind of probabilistic clustering model. Different from the k-means clustering model, GMM can give the probability that a sample belongs to a category. For example, the hybrid fund may belong to both the consumer category and the technology category. Therefore, this kind of clustering method is certainly practical and explanatory for fund clustering.

GMM is the combination of several Gaussian distributions. Assume there are  $k$  Gaussian distributions for the fund samples. Then the probability density function of the sample is shown as the following equation:

$$p(x) = \sum_{i=1}^k \alpha_i \cdot p\left(x|\mu_i, \sum_i\right), \quad (\text{B.1})$$

where  $x$  follows the mixed normal distribution  $N(\mu, \sum)$ ,  $\mu$  is the vector with means,  $\sum$  is the covariance matrix with a dimension of  $k \times k$ ,  $\mu_i$  and  $\sum_i$  are the mean and variance of the  $i^{\text{th}}$  Gaussian distribution,  $\alpha_i$  is the mixing coefficient of a single Gaussian distribution,  $\sum_{i=1}^k \alpha_i = 1$ , and  $\alpha_i \geq 0$ . The probability density function of a single Gaussian distribution is shown in the following equation:

$$p\left(x|\mu, \sum\right) = \frac{1}{(2\pi)^{n/2} |\sum|^{1/2}} e^{-((x-\mu)^T(x-\mu)/2\sum)}. \quad (\text{B.2})$$

As mentioned above, the parameters of the GMM are the mixing coefficient, the mean vector, and the covariance matrix  $\alpha, \mu, \sum$ . The maximum likelihood estimation method is adopted for parameter estimation. The maximum likelihood function is shown as equation (B.3). The analytical solution of the likelihood function cannot be obtained, so the idea of the maximum expectation (EM) algorithm is adopted to conduct parameter estimation.

$$L = \log \prod_{j=1}^n p(x) = \sum_{j=1}^n \log \left( \sum_{i=1}^k \alpha_i \cdot p\left(x|\mu_i, \sum_i\right) \right). \quad (\text{B.3})$$

Therefore, the probability that the JTH sample belongs to the ITH Gaussian distribution is shown in the following equation:

$$p(z_j = i|x_j) = \frac{\alpha_i \cdot p(x|\mu_i, \sum_i)}{\sum_{i=1}^k \alpha_i \cdot p(x|\mu_i, \sum_i)}. \quad (\text{B.4})$$

Based on the above analysis, assume the feature matrix of fund samples as  $(n$  and  $m)$ , where  $n$  is the fund sample number and  $m$  is the fund feature number. The GMM can be then carried out as follows:

*Step 6.* Initialize  $k$  Gaussian distributions with parameters  $\mu_i, \sum_i, \alpha_i$ .

*Step 7.* E step: calculate the probability that each sample belongs to each Gaussian distribution according to equation (B.4).

*Step 8.* M step: Update the mean vector  $\mu$  and the covariance matrix  $\Sigma$  of the Gaussian mixture distribution.

*Step 9.* Steps 2 and 3 are repeated until the increase in the loss function is less than a preset threshold or the maximum number of iterations is reached. The loss function is the likelihood function as shown in equation (B.3).

*Step 10.* Output the probability of each sample belonging to each Gaussian distribution and cluster the samples into the category with the largest probability.

We apply the Akaike information criterion (AIC) to evaluate the clustering results. AIC is a standard for evaluating the statistical model, which can be used to measure the model complexity and the goodness of fit. It can be obtained as shown in the following equation:

$$\text{AIC} = 2k - 2\ln(L), \quad (\text{B.5})$$

where  $k$  is a parameter, representing the model complexity, and  $L$  is the value of the likelihood function, representing the model goodness of fit. Generally speaking, the smaller the AIC value, the lower the model complexity, the higher the model goodness of fit, and the better the overall performance of the model.

## C. Deep Learning-Based Prediction Model

*C.1 Fully Connected Neural Network Layer.* The FCNN is generally the multilayer perceptron machine, including input layer, hidden layer, and output layer. In this study, we introduce the idea of the residual network (ResNet) into the FCNN, called residual FCNN (R-FCNN), namely adding a residual connection to the FCNN so as to alleviate the problem of gradient disappearance and gradient explosion in deep neural networks.

The ResNet is proposed in 2015 as shown in Figure 8 and equation (C.1). Assume the input is  $X$  and the output is  $F(X)$ . The  $F(X)$  usually includes operations such as convolution and activation. The idea of the ResNet is to add input  $X$  to the function of output  $F(X)$ , as shown in equation (C.1), which can be used to describe the nonlinear relationship between the input and output. Without using the new formula and theory, the residual connection just changes a new express so as to solve the problems of gradient disappearance and gradient explosion in deep network training.

$$x^{l+1} = F(x^l) + x^l. \quad (\text{C.1})$$

The architecture of the R-FCNN is shown in Figure 9. In this study, the R-FCNN is mainly used to extract the nonlinear correlation between input information and output information. The input is the semantic vector after embedding, and the output is the vector containing only one element, that is, the average fund value of a certain fund category in the next period.

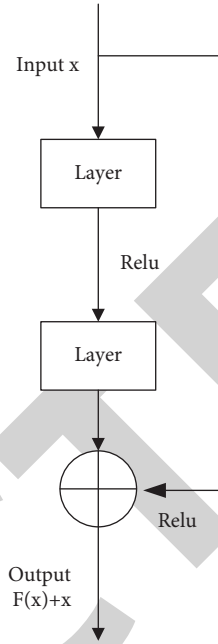


FIGURE 8: Diagram of residual connection.

*C.2 Long Short-Term Memory Layer.* A recurrent neural network (RNN) as shown in Figure 10 is a kind of powerful neural network that can deal with not only time series but also images. The input of RNN includes not only the current information but also the previous information. The historical information can be remembered by neurons and then passed forward through a feedforward neural network. The data flow is shown in equation (C.2), where  $\varnothing(\cdot)$  is the activation function,  $x_t$  is the input of the current time step,  $h_{t-1}$  is the saved historical information of the last time step, and  $W$ ,  $U$ , and  $V$  are the weight matrix.

$$\begin{aligned} h_t &= \varnothing(Wx_t + Uh_{t-1}), \\ y_t &= Vh_t. \end{aligned} \quad (\text{C.2})$$

However, when processing long sequence data, the RNN is likely to encounter the problem of gradient disappearance or gradient explosion, which makes RNN have only short-term memory, that is, RNN can only obtain the information of the near sequence when dealing with long sequence data but has no memory function for the earlier sequence, thus losing information. To solve this kind of problem, the LSTM structure is proposed by Hochreiter et al. LSTM is also a kind of recurrent neural network, which is mainly used to solve the problem that common RNN cannot remember long historical information. The data flow is shown in equation (C.3), where  $c$  is the memory cell matrix;  $\odot$  indicates the Hadamard product; the initiation of  $c$  and  $h$  is zero;  $\sigma$ ,  $\tanh$ , and sigmoid are activation function;  $W$ ,  $U$ , and  $V$  are weight matrices; and  $b$  is the bias vector.

A single LSTM cell is shown in Figure 11, which contains three gates, input gate ( $i_t$ ), forgetting gate ( $f_t$ ), and output gate ( $o_t$ ). They determine which information can be input, which information can be forgotten, and which information can be output, respectively. There is also a memory cell ( $c_t$ )

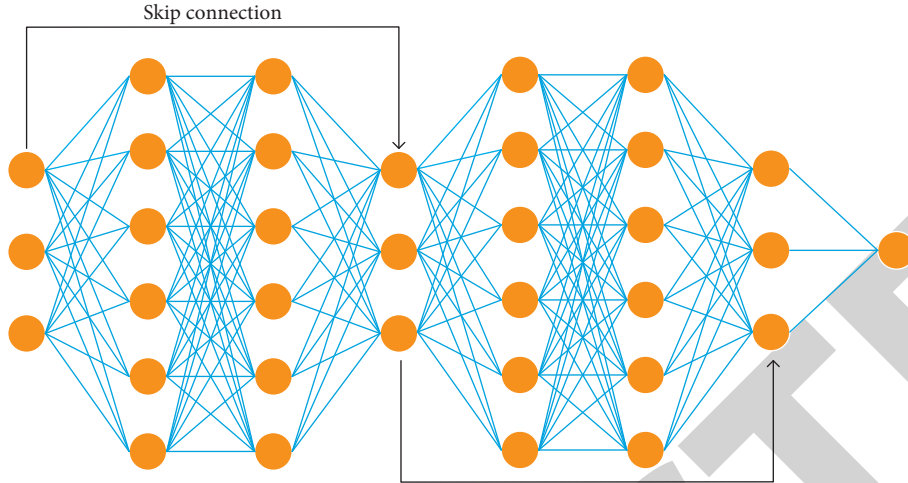


FIGURE 9: Diagram of the residual fully connected neural network layer.

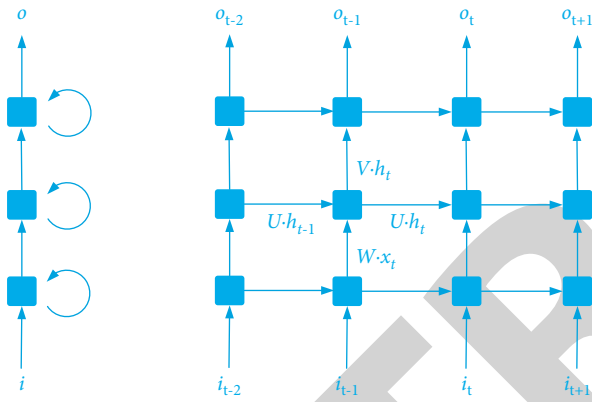


FIGURE 10: Diagram of recurrent neural network.

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + V_f c_{t-1} + b_f),$$

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + V_i c_{t-1} + b_i),$$

$$\tilde{c}_t = \tanh(W_c x_t + U_c h_{t-1} + b_c)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t,$$

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + V_o c_t + b_o) \quad (C.3)$$

$$h_t = o_t \odot \tanh(c_t),$$

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

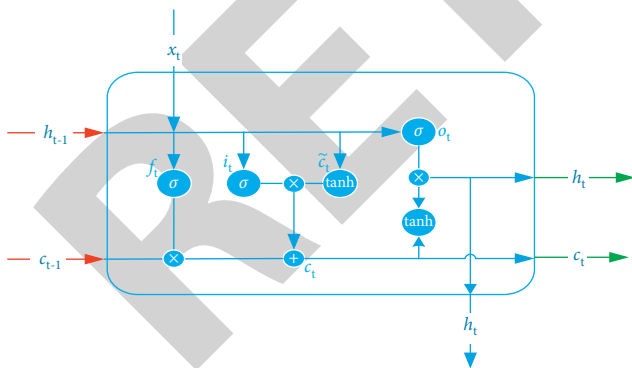


FIGURE 11: Diagram of LSTM cell.

that records the current state of the system and is controlled by three gates.

In this study, the LSTM is mainly applied to extract the temporal information in the fund data sequence. The input and output are the same as that of the R-FCNN, with the semantic vector after embedding as the input and the average fund value of a certain fund category in the next period as the output.

**C.3 Convolutional Neural Network Layer.** CNNs consist of one-, two-, and three-dimensional convolutional neural networks. The input data of the three operations correspond to different dimensions. The input data of the one-dimensional convolutional nerve needs to be one-dimensional, which is equivalent to the fully connected layer in the convolutional operation, as shown in Figure 12. The kernel in the convolution operation moves from left to right to get the final output. One-dimensional CNN is mainly used to extract the spatial correlation between input information and output information. Through the convolution operation of the multilayer convolutional kernel, the correlation between a single element and all other elements in the input information are effectively extracted.

Because the fund data is one-dimensional, we apply the one-dimensional CNN to extract the spatial correlations in the fund data sequence. The input is the semantic vector after embedding. The output of the CNN is

## Retraction

# Retracted: Generalized Zero-Adjusted Models to Predict Medical Expenditures

### Computational Intelligence and Neuroscience

Received 3 October 2023; Accepted 3 October 2023; Published 4 October 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] X. Xu, T. Ye, and D. Chu, "Generalized Zero-Adjusted Models to Predict Medical Expenditures," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 5874275, 18 pages, 2021.



## Research Article

# Generalized Zero-Adjusted Models to Predict Medical Expenditures

Xin Xu <sup>1</sup>, Tao Ye <sup>2</sup>, and Dongxiao Chu <sup>1</sup>

<sup>1</sup>School of Finance, Capital University of Economics and Business, Beijing 100070, China

<sup>2</sup>School of Banking and Finance, University of International Business and Economics, Beijing 100029, China

Correspondence should be addressed to Dongxiao Chu; chudongxiao@cueb.edu.cn

Received 22 October 2021; Accepted 17 November 2021; Published 13 December 2021

Academic Editor: Daqing Gong

Copyright © 2021 Xin Xu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In healthcare research, medical expenditure data for the elderly are typically semicontinuous and right-skewed, which involve a point mass at zero and may exhibit heteroscedasticity. The problem of a substantial proportion of zero values prevents traditional regression techniques based on the Gaussian, gamma, or inverse Gaussian distribution, which may lead to understanding the standard errors of the parameters and overestimating their significance. A common way to counter the problem is using zero-adjusted models. However, due to the right-skewness in the nonzeros' response, conventional zero-adjusted models such as zero-adjusted gamma, zero-adjusted Inverse Gaussian, and classic Tobit may not perform well. Here, we firstly generalize those three types of the conventional zero-adjusted model to solve the problem of right-skewness in health care. The generalized zero-adjusted models are very flexible and include the zero-adjusted Weibull, zero-adjusted gamma, zero-adjusted inverse Gaussian, and classic Tobit models as their special cases. Using the Chinese Longitudinal Healthy Longevity Survey, we find that, according to the AIC, SBC, and deviance criteria, the zero-adjusted generalized gamma model is the best one of these generalized models to predict the odds of zero cost accurately. In order to depict the predictors affecting the amount expenditure, we further discuss the situations where the mean, dispersion of a nonzero amount expenditure and model the probability of a zero amount of ZAGG in terms of predictor variables using suitable link functions, respectively. Our analysis shows that age, health, chronic diseases, household income, and residence are the main factors influencing the medical expenditure for the elderly, but the insurance is not significant. To the best of our knowledge, little study focused on these situations, and this is the first time.

## 1. Introduction

The ageing of the population is a universal law of the development of human society. According to the definition of the United Nations, if more than 10% of the total population of a country or region is seniors over 60 or over 7% are seniors over 65. The country or region has entered an ageing society. At present, most countries in the world, including the United States, the United Kingdom, and Japan, are about to experience the effects of population ageing. China has no exception and is also facing the complex ageing situation. The proportion of the population over 60 in China has increased from 10% in 1999 to 18% in 2019, nearly doubling in the next twenty years. The weakening of the physiological functions of the elderly will naturally increase their chances

of illness, leading to an increase in the demand for health services, which in turn brings about a large number of health and medical security problems for the elderly. Statistics from the China Health Commission showed that approximately 17% of the elderly consume nearly 70% of medical expenses. In addition, the ageing of the population will also inevitably bring about an increase in the life expectancy of the population. Experience has shown that chronic diseases are naturally associated with ageing. Therefore, the ageing population will result in a substantial increase in the prevalence of various chronic diseases. Ingmar et al. [1] discovered exceeding 60% of aged 65 and older had three or more coexisting chronic diseases in Germany. More than 180 million older adults in China suffer from chronic diseases, and the coexistence of multiple diseases is common.



According to the statistics of China's health and family planning, the medical expenditure of chronic diseases accounts for more than 70% of the health expenditure. However, due to the imperfections of China's existing medical security system, the out-of-pocket medical expense is relatively large, and medical security is insufficient [2, 3]. When the elderly's out-of-pocket medical expenditure reached or exceeded their ability to pay, some elderly people are not going to see the doctor after they become ill. Compared with other diseases, chronic diseases have the characteristics of insidious onset and long course, which will not only significantly increase the medical expenditure of the elderly [4, 5] but also cause some older adults to fail to receive medical treatment for minor illnesses. Therefore, accurate prediction of medical expenditure for the elderly will not only help the elderly arrange their consumption expenditure reasonably and improve their health status but also be advantageous to the country allocating medical resources more effectively.

Nevertheless, due to the different personal economic conditions of different elderly groups, they will have differences in medical expenditures after they fall ill. These phenomena will lead to a large number of zero consumption expenditures in the medical expenditure data of the elderly [6, 7], which also will result in right-skewed problems in the distribution of medical consumption data. Because of the point mass at zero and skewness, these problems can hardly be taken into account by traditional regression models such as Poisson, OLS, and gamma models. Ignoring these phenomena would lead to misspecified regression-based estimators and overestimated/underestimated effects. In order to predict more accurately, the medical expenses of the elderly new models need to be proposed. The aim of this paper is to propose a type of generalized zero-adjusted model to better fit the semicontinuous data, explore the influencing factors of elderly's medical expense, use this type of model to predict the amount of medical consumption of the elderly, and compare the results with conventional models.

The specific contributions of this paper include the following: (1) three types of generalized zero-adjusted models such as zero-adjusted generalized gamma model, zero-adjusted generalized inverse Gaussian model, and generalized zero-adjusted Tobit model for predicting the medical expenditure were proposed which included many traditional models and have not been used in health economic cost data modelling before. (2) Selected the best model zero-adjusted generalized gamma model according to different criteria and explored the marginal effects of predictors of medical expenditures. (3) Discovered the relationship between the dispersion of medical expenditure and explanatory variables due to the heterogeneity of variance.

The rest of this paper is organized as follows: detailed literature and related work are given in Section 2. Conventional zero-adjusted models and generalized zero-adjusted models are highlighted in Section 3. Numerous experimental results and comparisons of different models are suggested in Sections 4 and 5. Discussion is shown in Section 6. In the end, conclusions are summarized, and future research is presented in Section 7.

## 2. Related Works

There was much literature about modelling health care costs, although the health economist or health services researcher faced several difficulties. One type of published approaches for the medical expenditure involved modelling the cost using ordinary squares regression directly [8, 9]. Although the ordinary least squares model was used, this method was criticized because the distribution of strictly positive health expenditures was typically skewed, kurtotic (thick-tailed), and heteroskedastic, exhibiting a nonconstant variance that increased with expenditures [10]. These properties make the traditional approach, such as ordinary least squares (OLS) estimation biased and inefficient. Therefore, lots of work had been done on the problems of modelling medical expenditure. In order to solve the right-skewed of the medical expenditure data, Jones transformed the dependent variable using a log transformation to reduce the effect of extreme observations and right skewness and improved the goodness of fit [11]. Manning and Mullahy assumed the medical expenditures to be distributed to an exponential function of the explanatory variables and used log ordinary least squares and the gamma model with a log link to find a more robust alternative estimator than the OLS regression [12]. However, such transformations were likely to be problematic in heteroskedastic errors on the transformed scale [13, 14]. Alternately, there was a large and growing literature on using inherently nonlinear specifications to model medical expenditures, which benefited from estimating effects on the natural scale of costs. The generalized linear model (GLM) and exponential conditional mean models were considered. The generalized linear model was first proposed by Nelder and Wedderburn in the 70s last century and has been widely applied in many fields as once proposed [15]. The generalized linear model assumed that the dependent variable obeyed a type of exponential distribution family, which included many common distributions such as Poisson and normal and supposed the variance of the random error term was not required to be equal. Within the GLM family, it was usual to make assumptions about the functional form of the mean and variance of the distribution. Although the generalized linear model could effectively deal with the problem of heteroscedasticity, it perhaps failed to account explicitly for the issues of skewness and the fat tail, which had implications for the efficiency and robustness of estimators [16]. More flexible distributions for a greater range of estimated skewness and kurtosis coefficients were explored. Manning et al. proposed the generalized gamma models (GGM) to solve the problem of healthcare costs. The GGM model included important parametric distributions as nested and special cases, such as the gamma (GA) and log-normal (LN) distribution. Each model had been selected to model healthcare costs in lots of literature [14]. Because the GGMs was also a special limiting case of the generalized beta of the second kind (GB2), Jones investigated GB2 as part of a comparison of many different methods for modelling US healthcare costs. Mullahy [16] considered the use of the Singh–Maddala distribution (SM) in order to control the heavy right-hand tail of cost data, which was also nested within the GB2.

Where the censored approach to medical expenditure was concerned, the Tobit regression using a single distribution had been suggested as one of the methods to be used for modelling [17]. In Tobit regression, there was an assumption about the response variable based on a zero-truncated normal distribution. Obviously, the constant variance was assumed in this linear regression setting, and the response variable was right-skewed, which was inadequate for medical expenditure data. Therefore, the Gaussian assumption might not be suitable for fitting medical expenditure with the Tobit model. The censored Gamma regression was introduced to overcome the skewed nature of the response [18]. Unfortunately, the Tobit model could not also handle the excess zeros that was a phenomenon that there were more zeros than from the underlying distribution in the medical expenditure data. From the perspective of the data-generating process, semicontinuous medical expenditure data should be considered as arising from two different stochastic processes. Firstly, the patients might choose whether to see a doctor according to their health status, severity of illness, financial burden, and other reasons, which governed the occurrence of zeros. Secondly, the patients who enjoyed more medical services and higher income were likely to incur higher medical expenditures than those less inclined to use these services, which resulted in extremely asymmetry in the nonzero medical expenditures data. Therefore, a two-part mixture model was an ideal choice for dealing with such data, which separately model the probability of any medical services use and the level of expenditures conditional on use [19]. A large number of papers previously were explicitly devoted to changing the different distributions in the second process, and the binomial distribution or logistic regression model was used frequently in the first process. A log-normal distribution was often chosen to model the positive medical expenditure data [20]. However, many alternative distributions were used to relax the log-symmetry condition imposed by the log-normal distribution because the log-normal distribution was not enough to fit the right-skewed and heavy tail features in the data [21, 22]. To the best of our knowledge, there were many studies about the two-part mixture models used in medical healthcare, but the proposed approach had already been applied in many other fields. Heller et al. used the two-part model to predict the total claim count. The one part was the negative binomial distribution for modelling the claim counts, and the other was the inverse Gaussian for the claim amount that occurred. To estimate the total claim amount [23]. Chai et al. analyzed the semicontinuous arterial calcification scores by introducing a two-part skew log-normal [24, 25]. Liu et al. found that the generalized gamma model provided a superior fit in their analysis of daily alcohol consumption by comparing generalized gamma, log-skew-normal, and box-cox-transformed two-part models [26].

In recent years, there has been an increase in the use of Tweedie exponential family models to fit semicontinuous data [27]. The Tweedie family of distributions belongs to the exponential family with variance and has a compound Poisson-gamma interpretation with a probability mass at zero. The primary advantage of fitting such Tweedie models was to

avoid the two-part model of fitting the frequency and then the amount. It is a single distribution. Frees et al. predicted the insurance claim amount using Tweedie model [28]. Christoph F.K showed the better fit of the Tweedie model by comparing it with two-part models and Tobit model [29]. However, there was also another problem that the proposed Tweedie was not allowed to be fitted explicitly as a function of explanatory variables, according to Smyth and Jorgensen [27].

As an alternative, a recent study has perceived that a zero-adjusted regression model, which mixed discrete and continuous distribution. The discrete distribution of the zero adjusted regression model was represented by the Bernoulli distribution. In contrast, the continuous distribution can be represented by any continuous distribution with a positive range and right skewness. The zero-adjusted model could be regarded as a case of a two-part model. The zero-adjusted model focused more on the probability of zero value. When the probability of the observed zero value was much greater or less than the standard normal distribution, gamma, Weibull, and so on, the zero-adjusted model might be established. These could make the probability of zero occurrences predict more actually. Several applications of zero-adjusted gamma (ZAGA) and zero-adjusted inverse Gaussian (ZAIG) regression models could be found in insurance claims [30, 31]. Nevertheless, it appeared that there had been little work done in health economic cost data modelling before. This study attempted to use zero-adjusted models to predict medical expenditure. Throughout this paper, three types of generalized zero-adjusted models were presented, which comprised the classic Tobit model, ZAGA model and ZAIG model, and could improve the accuracy of the prediction. As far as we knew, there was almost little literature to study on those generalized zero-adjusted models, especially in the field of health care.

### 3. Methodology

**3.1. Spliced Distribution.** This study aims to use models to predict the medical expenditure for the elderly and discover the factors affecting the cost as accurately as possible. One way to dealing with excess zeros and positive skewness is to apply zero-adjusted models. The zero-adjusted model can be considered as a case of spliced distribution. Klugman et al. proposed a splicing method for creating new distributions [32], and it had been applied in modelling heavy tail for operational risk [33]. The density function of an  $n$ -component spliced distribution is defined as follows [32]:

$$f(x) = \begin{cases} a_1 f_1(x), & \text{if } x \in C_1, \\ a_2 f_2(x), & \text{if } x \in C_2, \\ \vdots & \\ a_n f_n(x), & \text{if } x \in C_n. \end{cases} \quad (1)$$

Here  $a_1, \dots, a_n$  are positive weights that add up to

$$a_1 + a_2 + \dots + a_n = 1. \quad (2)$$

The functions  $f_i(x) (i = 1, 2, \dots, n)$  are legitimate density functions with all probability on the interval  $C_i$ :

$\int_{C_i} f_i(x) dx = 1$ . The intervals  $C_i$  and  $C_j$  are mutually exclusive:  $C_i \cap C_j = \Phi, \forall i \neq j$ . The intervals of  $C_i$  are also sequentially ordered. That is to say,  $x < y$  if  $x \in C_i$  and  $y \in C_j$  for all  $i < j$ . There is an advantage of the spliced distribution allowing the inclusion of point mass distributions.

**3.2. Zero-Adjusted Model.** The zero-adjusted model can be regarded as a case of  $n$ -component spliced distributions when  $n$  equals 2. The first part has zero expenditure amounts, and the second part has nonzero expenditures, which are assumed to have a continuous distribution that accommodates heavy right-skewed. Let  $y_i$  be the expenditures of the  $i$ th older people,  $i = 1, \dots, n$ . The density function of zero-adjusted distribution may be written as follows:

$$f(y|x) = \begin{cases} \pi, & \text{if } y = 0, \\ (1 - \pi) \cdot g(y|x), & \text{if } y > 0. \end{cases} \quad (3)$$

where  $g(y)$  is the density of a continuous, right skewed distribution, and  $\pi$  is the probability of zero medical expenditure. The cumulative distribution of a zero-adjusted model (ZAM) can be expressed as

$$F(y|x) = \pi I_{\{y=0\}} + [\pi + (1 - \pi)G(y|x)] I_{\{y>0\}}, \quad (4)$$

where  $I(\cdot)$  is an indicator function.

**3.3. Discrete Part of ZAP.** Suppose the probability of an older person is distributed to the Bernoulli. Let  $\omega_i$  be a binary variable indicating the occurrence of the outcome for the older person with medical expenditure in one year and  $\pi_i$  be the probability of the positive medical expenditure, on person  $i$ .  $\pi_i$  may be a constant such as in equation (3) or be a random variable distributed as follows:

$$f(\pi_i|x) = \pi_i^{\omega_i} (1 - \pi_i)^{1 - \omega_i}, \quad \omega_i = 0, 1. \quad (5)$$

We consider the factors affecting the medical expenditure of the older person and incorporate covariates through the logit link function on  $\pi_i$ :

$$\log \frac{\pi_i}{1 - \pi_i} = \eta_i. \quad (6)$$

The predictor  $\eta_i$  is any form of a function related to factors, but generally is assumed to be a linear system  $\eta_i = \beta X_i$ ,  $\beta = (\beta_1, \dots, \beta_p)$ ,  $X_i$  is the vector of factors and  $\beta$  is the parameter. According to equation (6), we can predict the probability of medical expenditure for the elderly and determine the influencing factors of their medical decision-making.

**3.4. Continuous Part of the Zero-Adjusted Model.** Another advantage of spliced distributions is that they allow us to model different parts of a response variable with distributions. There are many candidate distributions for the nonzero heavy-tailed distribution modelling medical expenditures  $g(y|x)$ , such as the gamma, inverse Gaussian, log-normal (LN), Weibull (WEI), and log-skew-normal. In

this study, we considered two specifications of the spliced distributions. The first specification used the generalized gamma (GG) distribution, which includes the standard gamma, inverse gamma, Weibull, and log-normal distributions as special cases [21, 26]. Another was the generalized inverse Gaussian (GIG) distribution, which includes the inverse Gaussian as a special case and the gamma distribution and inverse gamma distributions as limiting cases [34, 35]. The Tobit distribution was also presented as a baseline comparison for the others, and we generalized the traditional Tobit model.

**3.4.1. Gamma Distribution (GA) and Inverse Gaussian Distribution (IG).** There was much work to deal with the problem of skewness and heteroscedasticity by transforming data. The data transformed seemed to be more homogeneous and symmetric. However, homoscedasticity was hardly achieved in fact, which resulted in biased estimation [36, 37]. Instead, we used gamma and inverse Gaussian distribution, which belonged to the generalized models taking into account heteroscedasticity and retaining the original dollar scale of the data. Furthermore, the gamma and inverse Gaussian models could accommodate skewness in the expenditures [38]. The GA model and IG model are included in the generalized linear model, which is mainly composed of three parts:

- (1) Systems part: the system part is a linear component which can be seemed like the traditional linear models similarly:

$$\eta_i = x_i' \beta, \quad (7)$$

where  $x_i$  is a column vector of covariates for observation  $i$ ,  $\beta$  is a column vector of the parameters, and  $\eta_i$  is a column vector of prediction  $i$ .

- (2) Link functions: the link functions  $g$  is often defined a monotonic and differential, which combines the prediction and the systems part and describes how the expected value of a response  $y_i$  is related to the linear predictors:

$$g(\mu_i|x) = x_i' \beta, \quad (8)$$

where  $g$  is often defined a log function.

- (3) Random parts: the response variables  $y_1, y_2, \dots, y_n$  are independent and distributed from an exponential family which implies there are a relationship between the variance and mean. The general form of exponential family is

$$f(y|\theta, \varphi) = \exp \left[ \frac{y\theta - b(\theta)}{a(\varphi)} + c(y, \varphi) \right], \quad (9)$$

where  $\theta$  is called the canonical parameter and represents the location, while  $\varphi$  is the dispersion parameter and represents the scale. Many distributions besides GA and IG models belong to the exponential distribution family, for example, normal, Weibull,

Poisson, negative binomial distributions, and so on. Because of skewness and heteroscedasticity of the outcome, the densities of the gamma distribution and inverse Gaussian distributions are

$$\text{gamma: } f(y|\mu, \sigma) = \frac{y^{1/\sigma^2-1} e^{-y/(\sigma^2\mu)}}{(\sigma^2\mu)^{1/\sigma^2} \Gamma(1/\sigma^2)}, \quad y > 0, \mu > 0, \sigma > 0. \quad (10)$$

Inverse Gaussian:

$$f(y|\mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2 y^3}} \exp\left[-\frac{1}{2\mu^2\sigma^2 y}(y - \mu)^2\right], \quad y, \mu, \sigma > 0. \quad (11)$$

Suppose the mean of gamma and inverse Gaussian distribution is  $E(Y|x) = \mu$ . The variance of gamma distribution is  $\text{Var}(Y|x) = \mu^2\sigma^2 = (E(Y|x))^2\sigma^2$ , and  $E(Y^r|x) = \mu^r \sigma^{2r} \Gamma(1/\sigma^2 + r) / \Gamma(1/\sigma^2)$ , for  $r > -1/\sigma^2$ . The skewness of gamma distribution is  $2\sigma$ , and excess kurtosis is  $6\sigma^2$ . The gamma distribution is appropriate for positively skewed data. At the same time, the variance of inverse Gaussian distribution is  $\text{Var}(Y|x) = \sigma^2\mu^3$ , and the skewness of inverse Gaussian is  $3\mu^{1/2}\sigma$ , the excess kurtosis is  $15\mu\sigma^2$ . The inverse Gaussian distribution is also appropriate for highly positively skewed data. We can see that the variance of response is a function of its mean. Note that the variance function for the inverse Gaussian GLM increases more rapidly with the mean than the gamma GLM.

**3.4.2. Generalized Gamma Distribution (GG).** Although the standard gamma model was fairly robust when we analyzed the positive medical expenditures([39]), it was inefficient when the data were heteroskedastic and heavily right-skewed([14]). The generalized gamma was available among other continuous distributions handling values only on the positive values. The density of the generalized gamma probability distribution is parameterized as a function of  $\kappa, \mu, \sigma$  and is given by [14, 21]

$$f(y; k, \mu, \sigma) = \frac{\theta^\theta}{\sigma y \sqrt{\theta} \Gamma(\theta)} \exp[z\sqrt{\theta} - \mu], \quad y \geq 0, \quad (12)$$

where  $\theta = |\kappa|^{-2}$ ,  $z = \text{sign}(\kappa)\{\ln(y) - \mu\}/\sigma$  and  $\mu = \theta \exp(|\kappa|z)^5$ . Because  $dz = (1/\sigma y)dy$ , equation (12) could be interpreted as the standard normal distribution ( $z$ ) scale for log-transformed  $y$ . If  $Y$  is a random variable distributed to density (12), then its mean was given by

$$E(Y|x) = \exp\left\{\mu + \frac{\sigma \log(\kappa^2)}{\kappa} + \log\left[\Gamma\left(\frac{1}{\kappa^2} + \frac{\sigma}{\kappa}\right)\right] - \log\left[\Gamma\left(\frac{1}{\kappa^2}\right)\right]\right\}. \quad (13)$$

The other moments of the generalized gamma distribution were  $m$ -th moment  $= E(Y^m|x) = [\exp(\mu) \cdot \kappa^{2\sigma/\kappa}]^m \{\Gamma[(1/\kappa^2) + (m\sigma/\kappa)]/\Gamma(1/\kappa^2)\}$

And, the variance was

$$\begin{aligned} \text{variance} &= E(Y^2|x) - E^2(Y|x) \\ &= [\exp(\mu) \cdot \kappa^{2\sigma/\kappa}] \\ &\quad \left\{ \left[ \frac{\Gamma((1/\kappa^2) + (2\sigma/\kappa))}{\Gamma(1/\kappa^2)} \right] - \left[ \frac{\Gamma((1/\kappa^2) + (\sigma/\kappa))}{\Gamma(1/\kappa^2)} \right]^2 \right\}. \end{aligned} \quad (14)$$

The standard gamma, inverse gamma, Weibull, and log-normal distributions were special cases of the generalized gamma distribution. For example, the generalized gamma distribution density reduced to a standard gamma distribution when the shape parameter  $\theta = \kappa^{-2}$  and the scale parameter  $\nu = \kappa^2 \exp(\mu)$ , i.e., the density follows as  $f(y; \nu, \theta) = (1/\nu^\theta \Gamma(\theta)) y^{\theta-1} \exp(-y/\nu)$ , and the mean was  $\exp(\mu)$ , the variance was  $\kappa^2 \exp(2\mu)$ . Let  $\kappa = -\sigma$ ,  $\sigma > 0$ , and the inverse gamma distribution was also obtained. The generalized gamma distribution reduced to an inverse gamma distribution defined as Robert [40], as follows.

$f(y; \varepsilon, \theta) = \varepsilon^\theta / \Gamma(\theta) (1/y)^{\theta+1} \exp(-\varepsilon/y)$ , where  $\varepsilon = \theta \exp(\mu)$ . When the parameter  $\kappa$  in equation (12) was fixed at a special value, for example,  $\kappa = 1$ , density (12) reduces to the probability density function of a Weibull distribution. In addition, if the parameter  $\kappa \rightarrow 0$ , density (12) reduced to the lognormal distribution, i.e.,  $f(y; \mu, \sigma) = 1/\sigma y \sqrt{2\pi} \exp\{-(\log(y) - \mu)^2/2\sigma^2\}$ .

**3.4.3. Generalized Inverse Gaussian Distribution (GIG).** We introduced the GIG distribution because the GIG was right-skewed, single-peaked distribution, and had a broader range of shapes. The standard gamma was a case of the GIG. Thus, the GIG could be a more flexible alternative to the standard version of the gamma [39]. The probability density function of the model was parameterized in terms of its mean, dispersion, and shape parameters. The parameterization of the generalized Inverse Gaussian distribution, denoted by  $GIG(\mu, \sigma, \nu)$ , was given by

$$f(y; \mu, \sigma, \nu) = \left(\frac{b}{\mu}\right)^\nu \left[\frac{y^{\nu-1}}{2K_\nu(\sigma^{-2})}\right] \exp\left[-\frac{1}{2\sigma^2}\left(\frac{by}{\mu} + \frac{\mu}{by}\right)\right], \quad (15)$$

for  $y > 0$ , where  $\mu > 0$ ,  $\sigma > 0$ , and  $-\infty < \nu < \infty$ . In the above equation (15),  $b = [K_{\nu+1}(\sigma^{-2})][K_\nu(\sigma^{-2})]^{-1}$  and  $K_\lambda(t)$  was a modified Bessel function of the second kind [40]:  $K_\lambda(t) = 1/2 \int_0^\infty x^{\lambda-1} \exp\{-1/2t(x+x^{-1})\} dx$ . With this parameterization, the mean  $E[Y|x] = \mu$  and variance  $\text{Var}[Y|x] = \mu^2 [2\sigma^2/b(\nu+1) + 1/b^2 - 1]$ . The skewness of GIG is

$$\begin{aligned} \text{skewness} &= \mu^3 \frac{[2 - 6\sigma^2/b(\nu+1)(\nu+2)\sigma^4 - 2/b^2 + 2\sigma^2/b^3(\nu+2)]}{[\text{Var}(y)]^{1.5}}, \\ \text{excess kurtosis} &= \mu^4 \frac{\left\{ \begin{aligned} &-6 + 24\sigma^2/b(\nu+1) + 4/b^2 [2 - \sigma^4(\nu+1)(7\nu+11)] + 4\sigma^2/b^3 [2\sigma^4(\nu+1)(\nu+2)(\nu+3) - 4\nu - 5] + \\ &1/b^4 [4\sigma^4(\nu+2)(\nu+3) - 2] \end{aligned} \right\}}{[\text{Var}(y)]^2}. \end{aligned} \quad (16)$$

Unlike the majority of models for insurance losses, our general approach could determine the distribution of each risk class based not only on the mean parameter, which was traditionally modelled in terms of covariates but also by using regressors on the dispersion and shape parameters, which described the shape of the GIG distribution. This could be regarded as a very useful property. Additionally, the GIG was a very wide family which included many well-known distributions depending on the estimated values of the dispersion and shape parameters which were modelled as functions of risk factors as was well known. For example, as could be seen,  $\text{GIG}(\mu, \sigma, -0.5) = \text{IG}(\mu, \sigma\mu^{-1/2})$ . Therefore, the gamma was a special case of GIG when  $b = 1$  and  $\nu = -1/2$ . According to Jorgensen [35],  $K_\lambda(t) \sim \Gamma(\lambda)2^{\lambda-1}t^{-\lambda}$  as  $t \rightarrow 0$ , for all  $\lambda > 0$ . And, when  $\sigma \rightarrow \infty$ ,  $\text{GIG}(\mu, \sigma, \nu)$  had limiting distribution  $GA(\mu, \nu^{-1/2})$  for all  $\nu > 0$ .

**3.4.4. Tobit Distribution.** Tobit model was first introduced to model dependent variables with a large fraction of zeros by Tobin [17]. The classic Tobit model assumed that the response was continuous, censored, and normally distributed underlying latent dependent variable  $y^*$ . We were interested in designing the latent variable  $y^*$  as a linear regression model:

$$\begin{aligned} y^* &= x_i'\beta + \varepsilon_i, \quad i = 1, \dots, n \\ y_i &= \begin{cases} y^*, & \text{if } y^* \leq L, \\ L, & \text{if } y^* > L, \end{cases} \end{aligned} \quad (17)$$

where  $\varepsilon \sim N(0, \sigma^2)$ ,  $x_i$  is an exogenous and observable explanatory variable. Specifically, if the latent variable  $y^*$  values equal to zero are censored, such as the medical expenditure for the elderly,  $L$  became zero. Then, the probability of a censored observation sample was

$$\Pr(y^* \leq L) = \Pr(x_i'\beta + \varepsilon \leq L) = \Phi\left[\frac{(L - x_i'\beta)}{\sigma}\right], \quad (18)$$

where  $\Phi(\cdot)$  was the standard normal cumulative distribution. We could present the truncated expected value of the noncensored observation  $y_i$

$$E(y_i | x_i, y_i > L) = x_i'\beta + \sigma \frac{\phi[(x_i'\beta - L)/\sigma]}{\Phi[(L - x_i'\beta)/\sigma]}, \quad (19)$$

where  $\phi(\cdot)$  was the density function of standard normal distribution. The classic Tobit model was appropriate when the response had two properties: one was that the error  $\varepsilon$

was a normal distribution, and the other was that the negative values of response were censored at  $L$ .

**3.4.5. New Type of Generalized Tobit Distribution.** The classic Tobit model was extremely sensitive to its underlying assumptions of normality and homoscedasticity. Therefore, the classic Tobit model should never be fit unless the data were truly normal and censored distribution. However, these were hardly met in real data [41, 42]. Many researchers claimed that a large mass at zero was censored observations when they were not censored, especially for health expenditure data. We provided another generalized Tobit model which was different from the generalized Tobit selection model by Heckman [43]. The Heckman selection model was considered as a generalized Tobit model and mainly connected the two latent outcomes by inverse mills ratio. However, the response variable was assumed to be normally distributed yet. In this paper, we mainly generalized the part where the latent  $y^*$  greater than zero. We selected the *student t* family model in order to compare the classic Tobit model. The *Student t* family model was introduced by Lange et al. [43] and was defined by assuming that  $Y = \mu + \sigma T$ , where  $T \sim t_\nu$  had a standard *t* distribution with  $\nu$  degrees of freedom. In this study, the PDF of *Student t* family distribution was given by

$$f_Y(y; \mu, \sigma, \nu) = \frac{1}{\sigma B(1/2, \nu/2)\nu^{1/2}} \left[ 1 + \frac{(y - \mu)^2}{\sigma^2 \nu} \right]^{-(\nu+1)/2}, \quad (20)$$

for  $-\infty < y < +\infty$ , where  $-\infty < \mu < +\infty$ ,  $\sigma > 0$ ,  $\nu > 0$ , and  $B(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$  is the beta function. Note that  $T = (Y - \mu)/\sigma$  had a standard *t* distribution with  $\nu$  degrees of freedom. It was obvious that the *t* distribution had higher kurtosis than the normal distribution and more suitable for modelling leptokurtic data [43]. The excess kurtosis of *t* distribution was  $6/\nu - 4$  (if  $\nu > 4$ ).

**3.4.6. Modelling the Probability of Zero Expenditures and Expected Nonzero Expenditures in Terms of Explanatory Variables.** We focused on the factors influencing the medical decision and the number of medical expenditures but paid attention to the accuracy of prediction. The mean  $\mu$  of regression model determined the number of medical expenditures and the dispersion  $\sigma$  could affect the accuracy

of prediction. Therefore, we would consider the zero-adjusted regression modes in different cases.

Case 1: when the probability of not seeing a doctor was constant, model (6) degenerated to

$$P(y = 0|x) = \pi, \quad (21)$$

$$\log(\mu) = \eta_1 = X_1^T \beta_1 + \varepsilon, \quad (22)$$

where the dispersion  $\sigma$  was considered as a constant and not affected by the predictor variables,  $X$  was the vector of predictors,  $\beta$  was the vector of parameters, and  $\varepsilon$  was the error. We used the log link function according to the exponential family [38].

Case 2: zero amount was not constant, which was affected by predictor variables, and the dispersion  $\sigma$  was still a constant. Then, the probability of a zero medical expenditure and mean is shown as

$$\logit(\pi) = \eta_2 = X_2^T \beta_2 + \varepsilon, \quad (23)$$

$$\log(\mu) = \eta_3 = X_3^T \beta_3 + \varepsilon, \quad (24)$$

where  $X$ ,  $\beta$ , and  $\varepsilon$  were the same as in equation (21) and (22). Theoretically, the factors that affected the decision-making equation (23) and the amount equation (24) were different. However, many studies assumed that they were the same, that was  $X_2 = X_3$ .

Case 3: the mean  $\mu$ , dispersion  $\sigma$ , and the probability of a zero amount  $\pi$  included in zero-adjusted model were all influenced by the predictors, which were modelled in terms of predictor variables using suitable link functions:

$$\logit(\pi) = \eta_4 = X_4^T \beta_4 + \varepsilon, \quad (25)$$

$$\log(\mu) = \eta_5 = X_5^T \beta_5 + \varepsilon, \quad (26)$$

$$\log(\sigma) = \eta_6 = X_6^T \beta_6 + \varepsilon, \quad (27)$$

where we could choose the same or different predictors  $X$  in equations (25)–(27). There had been much literature studying case 1 and case 2, and almost little to discuss case 3 to our best knowledge. In this study, we would analyze the three cases and compare their results.

**3.4.7. Maximum Likelihood Estimation.** According to the zero-adjusted model, given  $n$  independent observations  $y_i$  for  $i = 1, 2, \dots, n$ , the likelihood function was given by

$$L = L(\psi|y) = \prod_{i=1}^n f(y_i|x) = \prod_{i=1}^n \left[ \sum_{k=1}^2 \pi_k f_k(y_i|x) \right], \quad (28)$$

where  $\psi = (\theta, \pi)$ ,  $y = (y_1, y_2, \dots, y_n)^T$  and  $f_k(y_i) = f_k(y_i|\theta_k)$ . The log-likelihood function was given by

$$l = l(\psi|y) = \sum_{i=1}^n \log[\pi f_1(y_i|x) + (1 - \pi) f_2(y_i|x)]. \quad (29)$$

We wished to maximize the log-likelihood  $l$  concerning  $\theta$  and  $\pi$ . Nevertheless, the problem was that the logarithm of

the second summation in equation (29) made the solution difficult. In this paper, we used an algorithm provided by Rigby and Stasinopoulos [44] and was based on penalized likelihood estimation.

### 3.5. Model Validation and Verification

**3.5.1. Graphic Verification.** There were some approaches used for verification and selecting the best model among those models after fitting statistical probability models on the medical expenditures data, which mainly included two types of procedures: the graphical and the numerical approaches [45]. The graphical methods were used to verify whether the model described the systematic part and the independence of the normalized quantile residuals and their normality. In this study, we could obtain the mean, variance, skewness, and kurtosis to check the independence of the normalized quantile residuals and their normality by inspecting the residual versus fitted value plots, residual density plots, and Q-Q plots [44]. To assess the goodness-of-fit of the model, Akaike's information criterion (AIC) [46] and the Schwarz Bayesian criterion (SBC) [47] were considered as the numerical methods for validating and selecting the best model among the verified models. In addition, one goal of this study is to estimate the expected medical cost for individuals ( $\hat{y}_i = E[y_i|x_i]$ ). The mean prediction error can be thought as measuring the bias between the predicted outcome and the true response, which is often measured by the mean squared error (MSE).

**3.5.2. Information Criteria.** To compare the models and select the best one among the fitted models, we used the AIC, SBC, and the global deviance criteria. The AIC is computed based on the Kullback–Leibler distance in information theory, and the SBC is based on the integrated likelihood of Bayesian theory, which both impose the appropriate penalty on the average of the log-likelihood of models estimated given the number of coefficients estimated. A model with the lowest AIC and SBC values will be selected probably. The AIC is given as follows:

$$\text{AIC} = -2 \log(L) + 2p, \quad (30)$$

where  $L$  is the likelihood and  $p$  is the number of parameters in the model. The SBC is defined as:

$$\text{BIC} = -2 \log(L) + 2p \log(n), \quad (31)$$

where  $L$  and  $p$  are the same as in AIC and  $n$  is the sample size. As suggested by Rigby and Stasinopoulos [44] for parametric GAMLSS models, each model could be assessed by its fitted global deviance (GD) given by

$$\text{GD} = -2l(\hat{\theta}), \quad (32)$$

where  $l(\hat{\theta}) = \sum_{i=1}^n l(\theta_i)$ .

**3.5.3. Bias and Accuracy.** The bias measures the average deviation of the predicted value  $f(x)$  from the true value



$f(x)$  in a large number of the repeated sampling processes. The bias is often defined as followed given X:

$$\text{Bias}[\hat{f}(x)] = E[\hat{f}(x)] - f(x). \quad (33)$$

MSE can be thought of as measuring the bias of predictions and be defined as

$$\text{MSE}[\hat{f}(x)] = E[y - \hat{f}(x)]^2. \quad (34)$$

We can prove the MSE is minimized when  $\hat{f}(x) = E[y|x]$ . MSE is obtained through the sample value  $\text{MSE} = 1/N \sum_{i=1}^N (\hat{y}_i - y_i)^2$ , where  $\hat{y}$  denotes the estimated and  $y$  is the true value and  $N$  is the sample size. The MSE is an unbiased estimator of deviation.

**3.5.4. Out-of-Sample Analysis.** There is the last step to examine the appropriateness of the estimated models and the generalization ability of the model. We applied the bootstrap procedure to investigate how the results of our statistical analysis would be generalized to another data set. Given a data set  $D$  containing  $m$  samples, we could sample it and generate another data set  $D'$ . A sample was randomly selected from the data set  $D$  and put into the data set  $D'$ , and then the sample was put back into the initial data set  $D$ , so that the sample might be still drawn next time. After repeating this process  $m$  times, we were able to get a data set  $D'$  consisting of  $m$  samples. It was obvious that some samples in the data set  $D$  would appear multiple times in  $D'$ , while other samples maybe would not appear. The probability that a sample was never selected in  $m$  sampling was  $(1 - 1/m)^m$ , and its limit was  $\lim_{m \rightarrow \infty} (1 - (1/m))^m = 1/e \approx 0.368$ . About 36.8% of the samples in the data set  $D$  by the bootstrap sampling method did not appear in the sampling data set  $D'$ . In this way, we could use  $D'$  as the training set and  $(D - D')$  as the test set. In practical application, 1/3 of the sample size was generally selected as test set and 2/3 as the training set.

## 4. Empirical Analysis

**4.1. Data Description.** The aim of this paper was to discover the factors affecting the medical expenditures of the elderly in China and predict the amount using data from the Chinese Longitudinal Healthy Longevity Survey (CLHLS). This survey is a nationally representative panel study, containing observations on individuals aged 65 years or older covering more than half of the counties and cities from 23 provinces, cities, and autonomous regions in China. Since the start of the survey in 1998, it was repeated to follow the same group of the elderly every two or three years, which have been conducted eight waves until 2018. The survey includes questions about the health status, quality of life, medical care, and security needs of the elderly. We used data from the latest survey in 2018, which was a mixed cross-sectional data set collected from 1998 to 2018. In total, the sample consists of 15874 individuals. We finally selected 6832 samples after deleting the data with missing and no response. In what follows, we described variables retained for analysis. We began with the aimed response variable, the

medical expenditures, followed by other main independent variables such as income, health, and education.

**4.2. Description of Variables.** The distribution of medical expenditure for the elderly with the entire sample is shown in Figure 1. We could find that there were a large of zeros, and the histogram of the medical cost was right-skewed and heavily fat tail. From the empirical cumulative distribution plot (Figure 2), it could be seen that the medical expenditure data in the upper right part seriously deviates from the straight line. Therefore, the OLS regression model was not suitable for the data, and other transformed models must be considered. The histogram is shown in Figure 1 suggested a mixture of point distribution and a continuous distribution on the positive side. Consequently, a Tobit model may lead to biased inferences due to there being far more zero observations than expected under the Tobit formulation. The zero-adjusted models offer us a viable framework to deal adequately with the excess of zeros.

For this study, we were interested in revealing the factors affecting medical consumption behavior. In addition to the response variable, we included a set of explanatories in the regressions that turned out to affect the medical expenditure. Typical variables that emerged from the existing literature are age, sex, household income, marriage, and education [1–5]. We incorporated all these variables and added several other variables to the analysis that significantly improved the estimation of the zero-adjusted models. These variables describe the characteristics of the elderly, such as insurance, health status, action limited, individual education, residence, and heart disease or not.

The Andersen Behavioral Model of Health Service Use usually provided a framework for the study of hospitalization that outlined the three determinants: predisposing, enabling, and need factors [48]. In light of this, we evaluated the effects of health status and functional disabilities, as need factors and associated social-demographic factors, as predisposing and enabling factors, on hospitalization utilization. A complete list of the variables used and their descriptive statistics is presented in Table 1. We treated the variables medical expenditure, education, and household income as numerical. For the convenience of calculation, we divided the medical expenditure and the annual household income by 1000. All other variables were categorical and entered into the regressions as dummy variables.

The density distributions of  $f(y)$ , for nonzero medical expenditure are given in Figure 3. In this study, six right-skewed distributions were considered: the normal, student  $t$ , gamma, inverse Gaussian, generalized gamma and generalized inverse Gaussian distributions. The normal distribution was also presented as a benchmark comparison for the other, right-skewed distributions. All of the candidate distributions were subsequently fitted on a training set of a random 70% subsample. Figure 3 suggested that the normal distribution had the worst fit for the histogram of nonzero medical expenditure and other right-skewed distributions seemed to be fit better. The fitted values of the normal, inverse Gaussian, generalized gamma, and generalized

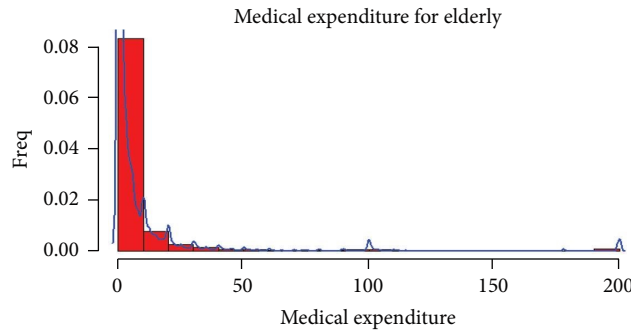


FIGURE 1: Histogram of medical expenditure.

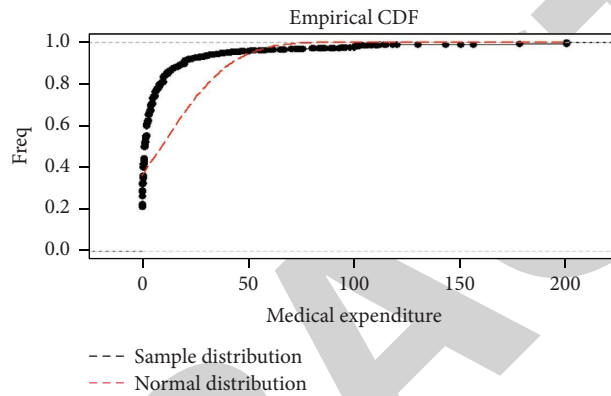


FIGURE 2: Cumulative distribution of medical expenditure.

TABLE 1: Descriptive statistics of the dependent and independent variables.

Variable	Description	Mean	S.D	Min	Max
Medical expenditure (numerical)	Medical expenditure/1000	9.034644	25.78056	0	199.998
Gender (categorical)	0 if female 1 if male	0.4428697	0.4967617	0	1
Age (categorical)	Age is divided four groups from low to high: lowest = 0 (reference); lower = 1; higher = 2; highest = 3	2.314582	1.129733	1	4
Health (categorical)	Very bad = 0(reference); bad = 1; so so = 2; good = 3; very good = 4	2.449445	0.9057199	0	4
Education (numerical)	Continuous	3.686622	4.413021	0	22
Actlim (categorical)	1 if limited in activities at least the last 6 months 0 if not limited (reference)	0.3261251	0.4688279	0	1
Married (categorical)	1 if married 0 if divorced, widowed, or never married (reference)	0.4584309	0.4983055	0	1
Insurance (categorical)	1 if participate in any insurance program 0 if no insurance (reference)	0.9211066	0.2695921	0	1
Residence (categorical)	1 if current interviewee lived in city (reference) 2 if lived in town 3 if lived in rural	2.17418	0.8024561	1	3
Heart_disease (categorical)	1 if suffering heart disease 0 if no heart disease (reference)	0.1891101	0.3916247	0	1
Household income (numerical)	Household income/1000	42.2063	36.4047	0	99

Note: medical expenses and annual household income are in thousands of yuan.

inverse Gaussian distributions underestimated the actual value at the lower points of medical expenditure. However, they showed better fit at other points. The fitted value of the gamma distribution overestimated the lower points.

Therefore, there seemed to be no obvious evidence to show which one was the best from the histogram. We must combine other statistical indicators to choose the best model, which was also done in the following section.

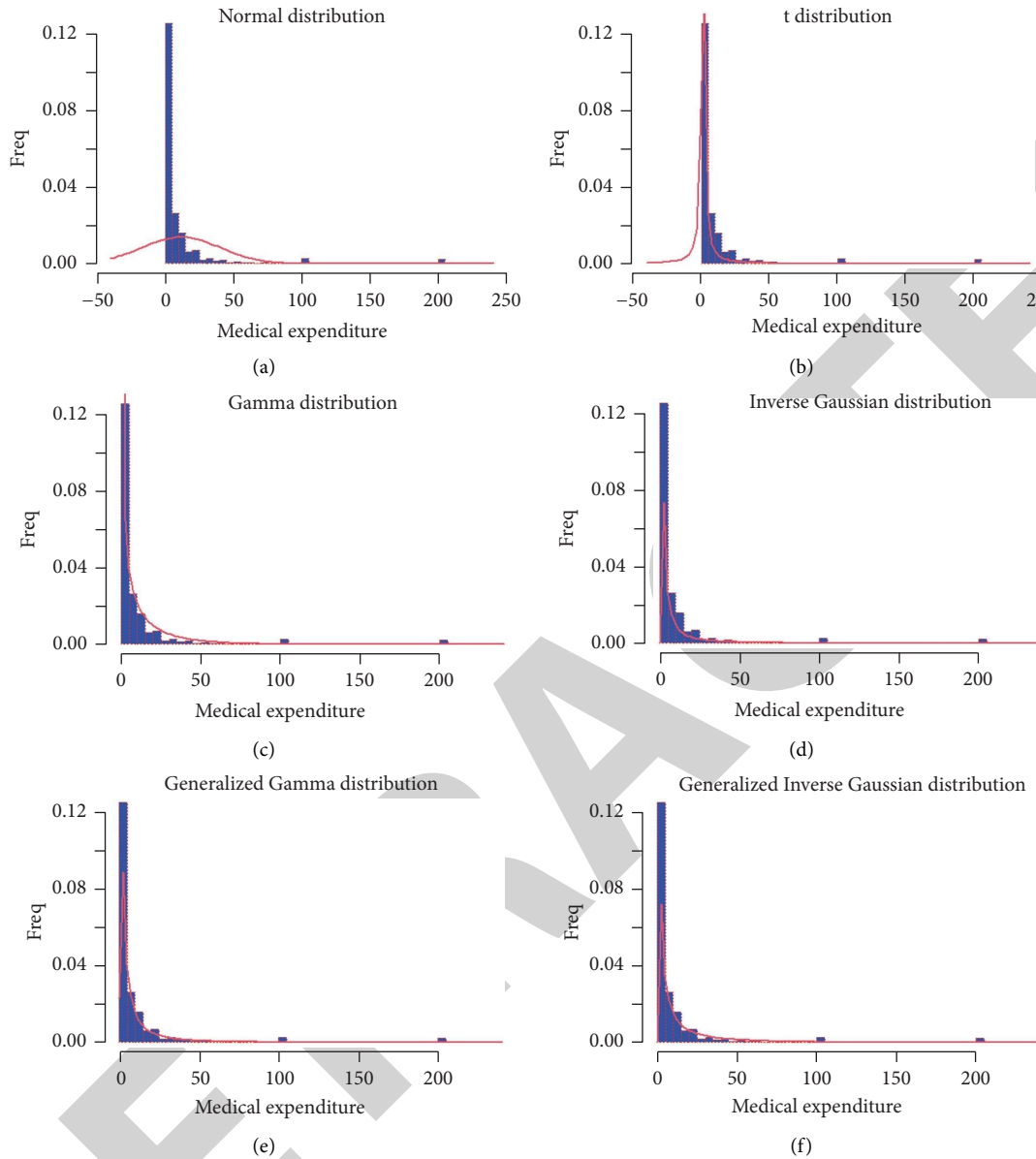


FIGURE 3: Candidate distributions for nonzero medical expenditure on the training set. (a) Normal distribution. (b)  $t$  distribution. (c) Gamma distribution. (d) Inverse Gaussian distribution. (e) Generalized gamma distribution. (f) Generalized inverse Gaussian distribution.

## 5. Results

According to the experience above, we chose 4872 samples as training data set. Table 2 listed the marginal effects estimation results obtained from the training for the models discussed above. We selected ten predictors according to the demand for health and health care of the Grossman model. The estimates of the Tobit model were quite different from others in the value range and sign, which had the largest values of AIC, SBC, and global deviance. This suggested that the Tobit model fitted the data very badly. The new generalized Tobit model and other zero-adjusted models were more similar. Many estimates shared the same sign and had comparable values, resulting in similar conclusions. In terms of the standard errors of the parameters, the errors of the Tobit model were

significantly higher than others. All zero-adjusted models had lower standard errors of the parameters. Furthermore, the AIC, SBC, and global deviance (GD) of the zero-adjusted generalized gamma model and zero-adjusted generalized inverse Gaussian model were obviously smaller than those of other models. However, the values of the zero-adjusted generalized gamma model were the smallest. The smaller these values are, the better the goodness-fit of the model is. Therefore, the ZAGG model was the best model we chose.

To assess model fit, we created the quantile residuals plots. If the models were adequate for the data, the residuals approximated a random sample from the standard normal distribution [38]. Figure 4 plots the normalized quantile residuals and shows much difference in the residuals among these models. The residuals of the new generalized Tobit

TABLE 2: Comparison of marginal effects of Tobit, generated Tobit, ZAGA, ZAIG, ZAGG, and ZAGIG models on CLHLS data.

	Tobit	Generalized Tobit	Discrete Binomial	Zero-adjusted models			
				Gamma	Continuous		
					Inverse Gaussian	Generalized gamma	Generalized inverse Gaussian
Intercept	17.30660 (4.62914)***	1.2211352 (0.3509476)***	-1.214247 (0.441916)***	3.4395522 (0.2634442)***	3.131784 (0.980320)***	1.3507565 (0.3023424)***	3.3181353 (0.2775036)***
Gender	-0.12960 (0.93239)	-0.1742161 (0.0706867)**	0.107757 (0.081745)	0.1653616 (0.0546928)***	0.423748 (0.149248)***	0.0757141 (0.0627683)	0.0798758 (0.0576116)
Age (lower)	0.37558 (1.17482)	0.0191777 (0.0890663)	-0.022572 (0.105939)	0.0503636 (0.0683660)	0.094369 (0.184435)	0.0347497 (0.0784604)	0.0439357 (0.0720145)
Age (higher)	-2.47297 (1.31612)*	-0.1307531 (0.0997782)	0.118543 (0.30750)	-0.0940778 (0.0771663)	-0.063349 (0.200707)	-0.1409848 (0.0885601)	-0.1377442 (0.0812844)*
Age (highest)	-7.43904 (1.48503)***	-0.4449018 (0.1125843)***	0.597306 (0.126676)***	-0.2929559 (0.0894135)***	-0.079564 (0.225999)	-0.4639923 (0.1026156)***	-0.3536105 (0.0941853)***
Health (bad)	-4.90168 (4.18061)	0.4308578 (0.3169431)	-0.466233 (0.413161)	-0.6101187 (0.2364191)***	-0.331020 (0.924567)	0.0096557 (0.2713270)	-0.3921744 (0.2490362)
Health (so so)	-7.88404 (4.09467)*	-0.1249660 (0.3104278)	-0.022842 (0.398988)	-0.7872073 (0.2319896)***	-0.678053 (0.907075)	-0.2875638 (0.2662435)	-0.6390054 (0.2443703)***
Health (good)	-11.95574 (4.11483)***	-0.4750543 (0.3119565)	0.263264 (0.399472)	-1.1046718 (0.2333300)***	-1.102426 (0.907025)	-0.7290380 (0.2677818)***	-0.9737652 (0.2457822)***
Health (very good)	-12.66309 (4.23632)***	-0.5853829 (0.3211667)*	0.291206 (0.407768)	-1.1434518 (0.2410609)***	-1.076392 (0.918506)	-0.8854640 (0.2766542)***	-1.1132889 (0.2539257)***
Education	0.12139 (0.12539)	0.0271759 (0.0095062)***	-0.016711 (0.011463)	-0.0053850 (0.0073003)	-0.007183 (0.021939)	0.0125635 (0.0083782)	-0.0071493 (0.0076899)
Actlim (limited)	6.86144 (0.97959)***	0.4495150 (0.0742651)***	-0.589516 (0.091230)**	0.2587853 (0.0565752)***	0.200809 (0.152316)	0.4854051 (0.0649287)***	0.2894195 (0.0595945)***
Household income	0.09582 (0.01294)***	0.0025677 (0.0009808)***	-0.002396 (0.001128)***	0.0081379 (0.0007603)***	0.010034 (0.002247)***	0.0043830 (0.0008726)***	0.0046224 (0.0008009)***
Marriage (married)	-1.64671 (1.04603)	0.1078140 (0.0793026)	0.005226 (0.092940)	-0.0363325 (0.0612575)	-0.008065 (0.160483)	0.0523691 (0.0703023)	0.0407834 (0.0645266)
Insurance (insured)	-5.71764 (1.53651)***	0.0433423 (0.1164868)	0.020459 (0.133651)	-0.4417009 (0.0904223)***	-0.296035 (0.260748)	-0.0773757 (0.1037734)	-0.2467232 (0.0952479)***
Residence (in town)	-6.38221 (1.26549)***	-0.4866704 (0.0959404)***	0.218619 (0.112552)*	-0.5132267 (0.0738881)**	-0.657867 (0.240905)***	-0.6409893 (0.0847979)***	-0.4931939 (0.0778313)***
Residence (in rural)	-5.01228 (1.25786)***	-0.5029975 (0.0953621)***	0.055642 (0.112483)	-0.5036751 (0.0733478)***	-0.693208 (0.238238)***	-0.8074203 (0.0841778)***	-0.5935885 (0.0772622)***
Heart_disease (suffered)	8.74595 (1.10292)***	0.9761225 (0.0836154)***	-1.139552 (0.132085)***	0.3834217 (0.0607861)***	0.503166 (0.198495)**	0.6753993 (0.0697614)***	0.4836672 (0.0640301)***
Sigma ( $\sigma$ )	3.35531 (0.01023)***	0.37450 (0.02461)***	—	0.398879 (0.009455)**	0.30921 (0.01156)***	0.53659 (0.01156)**	1.159849 (0.009905)***
AIC	37194.46	27567.15	—	26887.54	26951.03	25575.17	25817.39
SBC	37310.97	27690.13	—	27114.08	27177.57	25808.19	26050.41
GD	37158.46	27529.15	—	26817.54	26881.03	25503.17	25745.39
MSE	702.5649	751.1767	—	675.1498	744.6665	675.1498	671.3679

Note: ( ) indicates the standard errors of the parameters and the stars show the significance of the parameters: \*\*\*  $p < 0.01$ ; \*\*  $p < 0.05$ ; \*  $p < 0.1$ .

model showed bimodal kurtoses, ZAGA and classic Tobit models presented the aiguille characteristics, and the residuals of the ZAIG model appeared to right-skewed. The residuals of ZAGG and ZAGIG seemed to be similar, but the ZAGG model exhibited a better model fit from Figure 4.

We used worm plots to further study the residuals of these models. Worm plots of the residuals were introduced by van Buuren and Fredriks [49] to identify regions (intervals) of an explanatory variable within which the model does not adequately fit the data. These points in the worm plot, such as Figure 5, showed how far the ordered residuals were from their (approximate) expected values represented by the horizontal dotted line. The closer the points were to the horizontal line, the closer the

distribution of the residuals was to a standard normal distribution. Additionally, if the model was correct, we would expect approximately 95% of the points to lie between the two elliptic curves and 5% outside in Figure 5. A higher percentage of the points outside the two elliptic curves indicated that the fitted distribution of the model was inadequate to explain the response variable. The shape of the fitted curve to the points of the worm reflected different inadequacies in the model. A linear trend (positive or negative), quadratic shape (U or inverse U), or cubic shape (S shape) indicated a problem with the variance, skewness, or kurtosis of the residuals, respectively. This, in turn, highlighted a problem with the fitted distribution. Figure 5 shows that the fitted curves of the worm

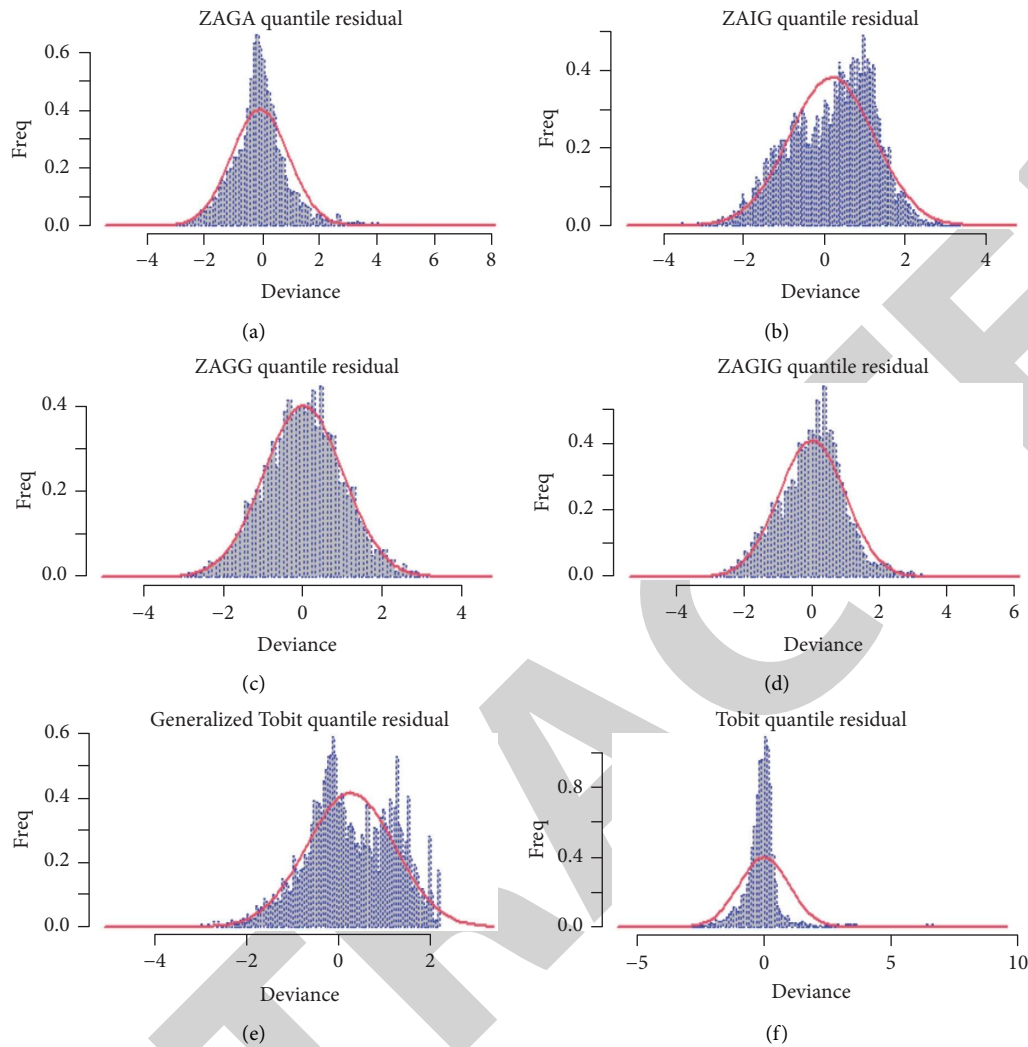


FIGURE 4: Quantile residual for regression models. (a) ZAGA quantile residual. (b) ZAIG quantile residual. (c) ZAGG quantile residual. (d) ZAGIG quantile residual. (e) Generalized Tobit quantile residual. (f) Tobit quantile residual.

for all models except the zero-adjusted generalized gamma model was S-shaped, which also suggested that the ZAGA model was generally a better fit again.

After the verification of 1960 samples in the test data set, two Tobit models and the ZAIG model had larger MSE values. ZAGG and ZAGIG models again produced very similar results. ZAGIG had the lowest MSE with a value of 671.3679. However, ZAGG was slightly higher with a value of 675.1498 than ZAGIG. Considering the goodness of fit, we chose the ZAGG model for further study.

We chose ZAGG models with different parameters  $\pi$ ,  $\sigma$ , and  $\nu$  and used the default log link function to discover which factors affecting the medical expenditure for the elderly. Table 3 shows the results of different parameters using 6832 data of the whole population. The predictors for the logarithm of average medical consumption shared almost the same sign and had similar values in the three models. Age, health, and chronic diseases were the main predictors influencing medical expenditure. With the increase of age, the medical expenditure of the elderly was decreasing. Part

of the reason might be that the elderly under 80 years old had a higher risk of serious diseases like cancer than the elderly over 80 years old. After experiencing this age stage, most of higher aged elderly were in good health. The elderly in good health had relatively less medical expenditure. As a chronic disease, heart disease significantly increased the medical expenses of the elderly. Compared with the elderly living in cities, the medical expenditure of the elderly living in urban and rural areas was relatively small, which might be related to the relative lack of medical resources in urban and rural areas of China. The higher the family income is, the more the medical expenditure of the elderly is. The predictor of medical insurance value was negative but not significant, which implied that medical insurance maybe reduced the medical expenditure of the elderly and released their financial burden. However, the effect was not obvious.

The scenario of the ZAGG(I) model: the proportion of  $p$  of zero medical expenditure and the scale parameter  $s$  relating to variance were both constant. Because the logit link function was used by default in the regression model, the

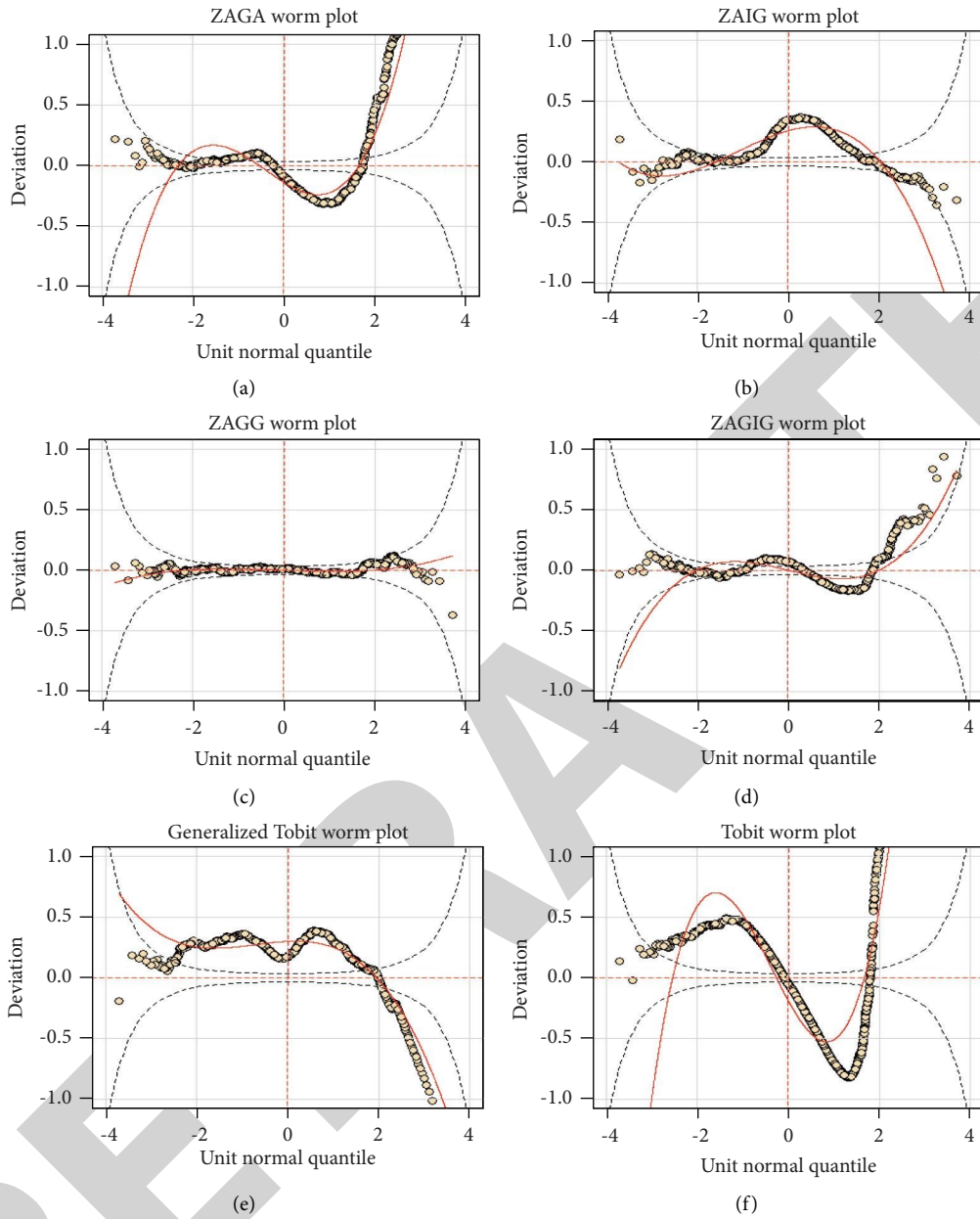


FIGURE 5: Worm plots for regression models. (a) ZAGA worm plot. (b) ZAIG worm plot. (c) ZAGG worm plot. (d) ZAGIG worm plot. (e) Generalized Tobit worm plot. (f) Tobit worm plot.

proportion of zero medical expenditure  $\pi$  was  $1/(1 + \exp(1.2984)) = 0.2144$ , which was very close to the proportion 0.2147 of zero cost in the population.

The scenario of the ZAGG (II) model: the occurrence of zero medical expenditure varied and was affected by the predictors. We found that some predictors for the medical decision shared different signs. For example, higher values of family income result in lower odds of zero medical expenditure. Perhaps, the elderly with high-income families are more likely to obtain medical resources, and the utilization of medical services was relatively high. The elderly with action limited were often in poor health, so their medical expenditures were more.

The scenario of the ZAGG (III) model: up to now, we had modelled the only  $\pi$  as a function of explanatory variables, but there were occasions in which the assumption of a constant scale parameter was not appropriate according to equation (14). On these occasions, modelling  $s$  as a function of explanatory variables could solve the problem. We could conclude from Figures 6–8 that almost all points of the worm plot failed inside the 95%-pointwise confidence intervals, indicating that the three models appeared to be adequate. Furthermore, the line shape with the negative slope in Figure 6 showed the variance was too low, and the fitted scale was too high. The s-shape with left bent down suggested the tails of fitted distribution was too light.



TABLE 3: Results from ZAGG models with different parameters  $\pi$ ,  $\sigma$ , and  $\nu$ .

	ZAGG(I)			ZAGG(II)			ZAGG(III)		
	$\log(\mu)$	$\text{logit}(\pi)$	$\log(\sigma)$	$\log(\mu)$	$\text{logit}(\pi)$	$\log(\sigma)$	$\log(\mu)$	$\text{logit}(\pi)$	$\log(\sigma)$
Intercept	1.3922 (0.2488)	-1.2984 (0.0295)***	0.5412 (0.0097)***	1.3922 (0.2489)***	-1.7032 (0.3400)***	0.5412 (0.0097)***	1.4065 (0.2409)***	-1.7032 (0.4000)***	0.4603 (0.1021)***
Gender	0.0928 (0.0528)			0.0929 (0.0628)	0.1177 (0.0686)*		0.0856 (0.0526)	0.1177 (0.0686)	0.0397 (0.0221)
Age (lower)	0.1085 (0.0656)			0.0347 (0.0785)	-0.0549 (0.0896)		0.0876 (0.0646)	-0.0549 (0.089)	0.0496 (0.0269)*
Age (higher)	-0.0960 (0.0740)			-0.1409 (0.0886)	0.1644 (0.0971)*		-0.1086 (0.0726)	0.1644 (0.0971)*	0.0238 (0.0302)
Age (highest)	-0.3056 (0.0857)***			-0.4640 (0.1026)***	0.5715 (0.1068)***		-0.3019 (0.0867)	0.5715 (0.1068)	0.0571 (0.0345)*
Health (bad)	-0.1828 (0.2206)			-0.0957 (0.2713)	0.0097 (0.3771)		-0.2121 (0.2107)	-0.0957 (0.3771)	-0.0001 (0.0911)
Health (so so)	-0.5136 (0.2161)**			-0.2876 (0.2662)	0.3518 (0.3658)		-0.5431 (0.2067)***	0.3517 (0.3658)	0.0182 (0.0891)
Health (good)	-0.9516 (0.2174)***			-0.7290 (0.2678)***	0.6387 (0.3660)*		-0.9822 (0.2083)***	0.6387 (0.3660)*	0.0412 (0.0894)
Health (very good)	-1.0126 (0.2249)***			-0.8855 (0.2767)***	0.7286 (0.3720)*		-1.0339 (0.2169)***	0.7286 (0.3720)*	0.0688 (0.0926)
Education	0.0183 (0.0070)***			0.0126 (0.0084)	-0.0096 (0.0096)		0.0199 (0.0070)***	-0.0095 (0.0096)	-0.0024 (0.0029)
Actlim	0.4699 (0.0540)***			0.4854 (0.0649)***	-0.5312 (0.0755)***		0.47731 (0.0536)***	-0.5312 (0.0755)***	-0.00853 (0.0222)
House hold income	0.0048 (0.0007)***			0.0044 (0.0009)***	-0.0031 (0.0010)***		0.0051 (0.0008)***	-0.0031 (0.001)***	0.0014 (0.0003)***
Marriage (married)	0.0393 (0.0590)			0.0524 (0.0703)	-0.0318 (0.0785)		0.0412 (0.0583)	-0.0318 (0.0785)	-0.0494 (0.0238)**
Insurance (insured)	-0.0746 (0.0867)			-0.0774 (0.1038)	0.1055 (0.1145)		-0.0646 (0.0876)	0.1055 (0.1145)	-0.0278 (0.0359)
Residence (in town)	-0.5439 (0.0712)***			-0.6410 (0.0848)***	0.2107 (0.0947)**		-0.5264 (0.0711)***	0.2107 (0.0947)**	-0.0178 (0.0299)
Residence (in rural)	-0.6579 (0.0702)***			-0.8074 (0.0842)***	0.0926 (0.0941)		-0.6408 (0.0712)***	0.0926 (0.0941)	0.0418 (0.0291)
Heart_disease (suffered)	0.6637 (0.0588)***			0.6754 (0.0698)***	-1.1524 (0.1123)***		0.65708 (0.0569)***	-1.1524 (0.1123)***	-0.0459 (0.0242)*
AIC	—			37235.64	36845.32		36819.29	36819.29	
SBC	—			37372.23	37091.18		37174.42	37174.42	
Dev	—			37195.64	36773.32		36715.29	36715.29	
K-S (D)	—			0.012776	0.012171		0.010868	0.010868	
(p value)	—			(0.2147)	(0.2636)		(0.3951)	(0.3951)	

Note: () indicates the standard errors of the parameters, and the stars show the significance of the parameters: \*\*\*  $p < 0.01$ ; \*\*  $p < 0.05$ ; \*  $p < 0.1$ .

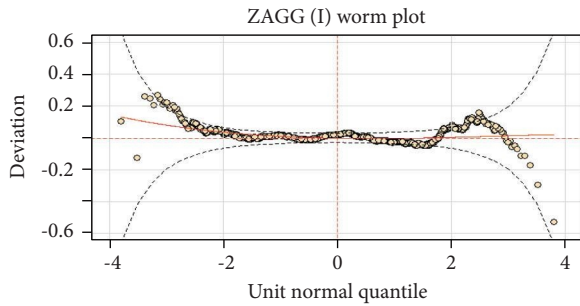


FIGURE 6: Worm plots for ZAGG(I).

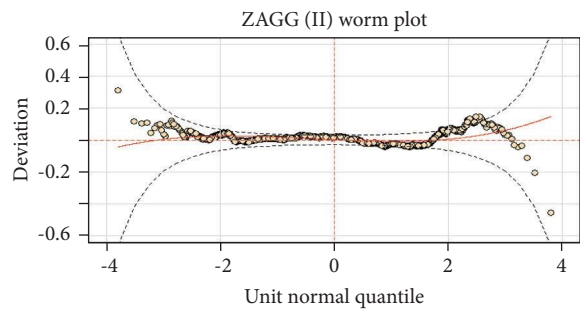


FIGURE 7: Worm plots for ZAGG(II).

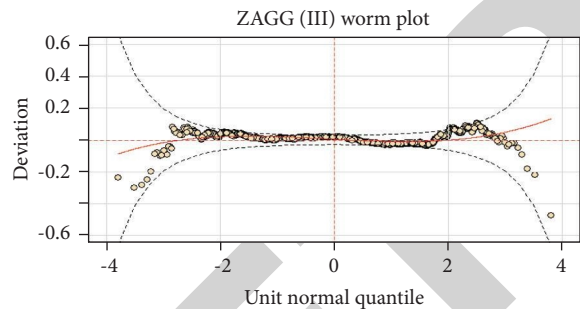


FIGURE 8: Worm plots for ZAGG(III).

Kolmogorov-Smirnov test is a nonparametric test method that can be used to compare the cumulative empirical distributions of two samples. The  $D_n$  ( $D_n = \sup_x |F_n(x) - F(x)|$ ) statistics is used to compare the maximum value of the difference between the empirical distributions of two samples. If this value is too large, we believe that the two distributions are different. Therefore, we used the two tailed Kolmogorov-Smirnov test to verify the consistency between ZAGG model and empirical distribution. The results are shown in the last row of Table 3. From the results, the  $p$  values were all greater than 0.05, which meant we could not reject the null hypothesis that there was almost no difference between the ZAGG model and empirical distribution.

## 6. Discussion

This paper explored and empirically validated zero-adjusted models with semiparametric formulation for estimating medical expenditures using CLHLS survey data. In reaction

to the limitations of conventional Tobit, zero-adjusted gamma, and zero-adjusted inverse Gaussian models, we generalized the three models to improve the accuracy of prediction and discover the factors affecting the elderly medical decision. The zero-adjusted generalized gamma model outperformed the zero-adjusted generalized Tobit and zero-adjusted generalized Inverse Gaussian model. Thus, the ZAGG mode provided an interesting alternative for modelling health care utilization expenditure data as it included many conventional models such as the zero-adjusted Weibull model, the zero-adjusted lognormal model, and the zero-adjusted gamma model. The ZAGG model included log-additive components for the mean and dispersion of medical expenditure given that expenditure occurs, as well as a logistic additive component for the probability of a zero expenditure. The model components were estimated independently and could be fitted with the same set of covariates. In this paper, we firstly chose ZAGG models with different parameters  $\pi$ ,  $\sigma$ , and  $\nu$  and used the default log link function to discover which factors affecting the medical expenditure for the elderly. There was much literature on the influence of factors on the parameter  $\pi$ , but there was almost no work discussing the influence of factors on the parameter  $\sigma$  and  $\nu$ . We found that some factors might affect the distribution shape and scale change of ZAGG model and then affected the accuracy of the model. These were also contributions of this paper.

Our empirical application had focused on the assessment of the predictive accuracy and the predictors affecting medical expenditure. We found that ZAGG and ZAGIG gave similar results. Moreover, ZAGG was appreciated for the fact that the generalization errors of ZAGIG were 671.36, which was less than that of ZAGG from mean square errors. However, the ZAGG model seemed to perform better in the aspects of global deviance, AIC and SBC. Whether one of the two models is superior to the other remains an open question, which needs to be determined according to different problems and situations. ZAGG and ZAGIG models, respectively, extended the ZAGA and ZAIG models, and many conventional zero-adjusted models are special forms of these generalized models. Moreover, both of these generalized models increased the difficulty in parameter estimation. For example, the standard errors of parameters were not reliable when the QR decomposition method was used, which could not solve the Hessian matrix. In this paper, we reported the QR-based standard errors using a likelihood-based confidence interval method introduced by Rigby and Stasinopoulos [44, 50, 51].

Although the ZAGG model is complex in application and calculation, it still has some advantages. One benefit of the ZAGG model is that the three components of the mixture model provide the analyst with a three-way interpretation by estimating the factors affecting the medical decision, the factors predicting the amount of the medical expenses, and the factors influencing the dispersion of the expenditure amount. The scale dispersion estimates can be used to provide more conservative estimates when the parameters were less robust. Another advantage of the ZAGG

model is that the regression method does not imply a “black box” approach for interpreting the effects of individual covariates. The interpretation of the marginal effect for the model is relatively explicit.

We surprisingly discovered that basic medical insurance had no significant effect on the medical expenditure of the elderly. The main reason was that the basic medical insurance had covered nearly 95% of the population in China up to now, leading to no obvious difference in the impact of medical insurance [2, 52–54]. The elderly with high-income families spent more on health care, which indicated a relatively unfair phenomenon that the poor subsidized the rich in the utilization of medical services in China. At the same time, the medical expenditure in urban and rural areas was relatively low, which also showed that the distribution of medical resources was not balanced.

Finally, attention should be paid to the limitations of our study. One limitation of this study was that it did not consider the causal relationship between the predictors and response because we were interested in predicting the amount of medical expenditure and unraveled the significant predictors influencing the expenditure. Possible solutions for this causal relationship were either to study only the impact of truly exogenous independent or to apply instrumental variable techniques. Another limitation was that the zero-adjusted models were seemed to be two-stage models, and there existed a variety of models in the continuous part. In this paper, we compared only a few models. Instead, other types of skewed distributions could be considered for further research. Finally, our study has used the two-stage model to predict the amount of medical expenditure. We treated the two parts as independent. However, there perhaps existed a relationship. There were further opportunities to develop potentially superior models by considering the correlation, such as copula function. Moreover, it should be noted that if the relationship were considered, the difficulty of parameter estimation would increase, and the effects of individual explanatory variables could not be interpreted conveniently.

## 7. Conclusions

In this paper, we have predicted the amount of medical expenses for the elderly and explored the marginal effect of the predictors in China, using CLHLS survey data. In reaction to the limitations of conventional Tobit and zero-adjusted models, we generalized these models. This allowed us to estimate the medical expenditure using more flexible models. The zero-adjusted generalized gamma model was the best to fit this data. We focused on the zero-adjusted generalized gamma regression model to reveal the significant factors influencing the medical amount. Several conclusions could be drawn from this work. The health status, family income, residence, and chronic diseases of the elderly significantly affect their medical expenditure, while the influence of basic medical insurance is not significant. We used a logistic model to discover the factors that affected the medical decision of the elderly. We found that the elderly in

the higher age group had the lower occurrence of zero medical amount, which indicated they were better health. In addition, this paper accurately estimated the proportion of the elderly with zero medical expenditure using a logit model. In the ZAGG model, we found that the scale dispersion was also affected by the explanatory variables, which could improve the robustness of the standard errors of parameters.

To the best of our knowledge, this is the first time that using the zero-adjusted generalized gamma model predicts the medical expenditure. The current approach appeared to be effective. However, some limitations merit attention, such as the causal relationship between the predictors and response and the correlations of the two parts in the zero-adjusted models. These limitations required further investigation in the near future.

## Data Availability

The data in this paper are obtained from the Chinese Longitudinal Healthy Longevity Survey (CLHLS) in 2018.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References

- [1] S. Ingmar, “Which chronic diseases and disease combinations are specific to multimorbidity in the elderly? Results of a claims data based cross-sectional study in Germany,” *BMC Public Health*, vol. 11, no. 1, pp. 1–9, 2011.
- [2] X. Li and W. Zhang, “The impacts of health insurance on health care utilization among the older people in China,” *Social Science & Medicine*, vol. 85, pp. 59–65, 2013.
- [3] K. Xu, D. B. Evans, G. Carrin, A. M. Aguilar-Rivera, P. Musgrove, and T. Evans, “Protecting households from catastrophic health spending,” *Health Affairs*, vol. 26, no. 4, pp. 972–983, 2007.
- [4] C. Yuyu and J. Ginger Zhe, “Does health insurance coverage lead to better health and educational outcomes? Evidence from rural China,” *Journal of Health Economics*, vol. 31, no. 1, pp. 1–14, 2012.
- [5] L. Cheng, H. Liu, Y. Zhang, K. Shen, and Y. Zeng, “The impact of health insurance on health outcomes and spending of the elderly: evidence from China’s new cooperative medical scheme,” *Health Economics*, vol. 24, no. 6, pp. 672–691, 2015.
- [6] Y. Min and A. Agresti, “Modeling nonnegative data with clumping at zero: a survey,” *Journal of the Iranian Statistical Society*, vol. 1, no. 1, pp. 1–2, 2002.
- [7] J. Aitchison, “On the distribution of a positive random variable having a discrete probability mass at the origin,” *Journal of the American Statistical Association*, vol. 50, no. 271, pp. 901–908, 1955.
- [8] S. Martin, *Developing a Person Based Resource Allocation Formula for General Practices in England*, University of York, New York, NY, USA, 2015.
- [9] D. Jennifer, “A person based formula for allocating commissioning funds to general practices in England: development of a statistical model,” *BMJ*, vol. 343, 2011.

- [10] D. K. Blough, C. W. Madden, and M. C. Hornbrook, "Modeling risk using generalized linear models," *Journal of Health Economics*, vol. 18, no. 2, pp. 153–171, 1999.
- [11] A. M. Jones, "Health econometrics," *Handbook of Health Economics*, vol. 1, 2000.
- [12] W. G. Manning and J. Mullahy, "Estimating log models: to transform or not to transform?" *Journal of Health Economics*, vol. 20, no. 4, pp. 461–494, 2001.
- [13] N. Duan, "Smearing estimate: a nonparametric retransformation method," *Journal of the American Statistical Association*, vol. 78, no. 383, 1983.
- [14] W. G. Manning, A. Basu, and J. Mullahy, "Generalized modeling approaches to risk adjustment of skewed outcomes data," *Journal of Health Economics*, vol. 24, no. 3, pp. 465–488, 2005.
- [15] J. A. Nelder and R. W. M. Wedderburn, "Generalized linear models," *Journal of the Royal Statistical Society: Series A*, vol. 135, no. 3, pp. 370–384, 1972.
- [16] J. Mullahy, "Econometric modeling of health care costs and expenditures: a survey of analytical issues and related policy considerations," *Medical Care*, vol. 47, no. 7, pp. S104–S108, 2009.
- [17] Tobin and James, "Estimation of relationships for limited dependent variables," *Econometrica*, vol. 26, no. 1, pp. 24–36, 1958.
- [18] F. Sigrist and W. A. Stahel, "Using the censored gamma distribution for modeling fractional response variables with an application to loss given default," *Papers*, vol. 41, no. 2, pp. 673–710, 2012.
- [19] N. Duan, W. G. Manning, C. N. Morris, and J. P. Newhouse, "A comparison of alternative models for the demand for medical care," *Journal of Business & Economic Statistics*, vol. 1, no. 2, p. 115, 1983.
- [20] J. G. Cragg, "Some statistical models for limited dependent variables with application to the demand for durable goods," *Econometrica*, vol. 39, no. 5, pp. 829–844, 1971.
- [21] L. Lei, "A flexible two-part random effects model for correlated medical costs," *Journal of Health Economics*, vol. 29, no. 1, pp. 110–123, 2010.
- [22] E. Mihram, "Parameter estimation for generalized gamma distribution," *Technometrics*, vol. 7, no. 3, pp. 349–358, 1965.
- [23] G. Z. Heller, D. Mikis Stasinopoulos, R. A. Rigby, and P. De Jong, "Mean and dispersion modelling for policy claims costs," *Scandinavian Actuarial Journal*, vol. 2007, no. 4, pp. 281–292, 2007.
- [24] H. S. Chai and K. R. Bailey, "Use of log-skew-normal distribution in analysis of continuous data with a discrete component at zero," *Statistics in Medicine*, vol. 27, no. 18, pp. 3643–3655, 2010.
- [25] A. Azzalini, "A class of distributions which includes the normal ones," *Scandinavian Journal of Statistics*, vol. 12, no. 2, pp. 171–178, 1986.
- [26] L. Lei, "Analyzing repeated measures semi-continuous data, with application to an alcohol dependence study," *Statistical Methods in Medical Research*, vol. 25, no. 1, pp. 133–152, 2012.
- [27] B. Jorgensen and M. C. P. D. Souza, "Fitting tweedie's compound Poisson model to insurance claims data: dispersion modelling," *Astin Bulletin*, vol. 32, no. 1, pp. 143–157, 2002.
- [28] F. Edward, L. Gee, and Y. Lu, "Multivariate frequency-severity regression models in insurance," *Risks*, vol. 4, no. 1, 2016.
- [29] C. F. Kurz and F. Christoph, "Tweedie distributions for fitting semicontinuous health care utilization cost data," *BMC Medical Research Methodology*, vol. 17, no. 1, p. 171, 2017.
- [30] G. Heller, D. Stasinopoulos, and R. Rigby, *The Zero-Adjusted Inverse Gaussian Distribution As A Model For Insurance Claims*, Statistical Modelling Society, Galway, Ireland, 2006.
- [31] A. B. Bortoluzzo and D. P. Claro, *Estimating Claim Size and Probability in the Auto-Insurance Industry: The Zero-Adjusted Inverse Gaussian (ZAIG) Distribution*, Statistical Modelling Society, Galway, Ireland, 2009.
- [32] S. A. Klugman, *Loss Models: From Data to Decisions*, Wiley, , Wiley, New York, NY, USA, 4th edition.
- [33] G. Peters and P. Shevchenko, *Advances in Heavy Tailed Risk Modeling: A Handbook of Operational Risk*, Wiley, New York, NY, USA, 2015.
- [34] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, Houghton Mifflin, Boston, MA, USA, 1970.
- [35] B. Jorgensen, *Statistical Properties of the Generalized Inverse Gaussian distribution*, Springer, Berlin, Germany, 1982.
- [36] W. G. Manning, "The logged dependent variable, heteroscedasticity, and the retransformation problem," *Journal of Health Economics*, vol. 17, no. 3, pp. 283–295, 1998.
- [37] J. Mullahy, "Much ado about two: reconsidering retransformation and the two-part model in health econometrics," *Journal of Health Economics*, vol. 17, no. 3, pp. 247–281, 1998.
- [38] P. M. McCullagh and J. Nelder, "Generalized linear models," *Applied Stats*, vol. 393, 2nd edition, 1989.
- [39] C.-W. Chou and W.-J. Huang, "On characterizations of the gamma and generalized inverse Gaussian distributions," *Statistics & Probability Letters*, vol. 69, no. 4, pp. 381–388, 2004.
- [40] C. P. Robert, "The bayesian choice. from decision-theoretic foundations to computational implementation," *The Bayesian Choice*, Springer, Berlin, Germany, 2nd edition, 2007.
- [41] M. Hurd, "Estimation in truncated samples when there is heteroscedasticity," *Journal of Econometrics*, 1979.
- [42] A. S. Goldberger, "Linear regression after selection," *Journal of Econometrics*, vol. 15, no. 3, pp. 357–366, 1981.
- [43] K. L. Lange, R. Little, and J. Taylor, "Robust statistical modeling using the T distribution," *Journal of the American Statistical Association*, vol. 84, no. 408, 1989.
- [44] R. A. Rigby and D. M. Stasinopoulos, "Generalized additive models for location, scale and shape," *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 54, no. 3, 2005.
- [45] A. Md, "Best-fit probability distributions and return periods for maximum monthly rainfall in Bangladesh," *Climate*, vol. 6, no. 1, p. 9, 2018.
- [46] H. Akaike, "Maximum likelihood identification of gaussian autoregressive moving average models," *Biometrika*, vol. 60, 1973.
- [47] G. E. Schwarz, "Estimating the dimension of a model," *Annals of Statistics*, vol. 6, no. 2, 1978.
- [48] R. M. Andersen, "Revisiting the behavioral model and access to medical care: does it matter?" *Journal of Health and Social Behavior*, vol. 36, no. 1, pp. 1–10, 1995.
- [49] S. v. Buuren and M. Fredriks, "Worm plot: a simple diagnostic device for modelling growth reference curves," *Statistics in Medicine*, vol. 20, no. 8, pp. 1259–1277, 2001.
- [50] B. Rigby and M. Stasinopoulos, *A Flexible Regression Approach using GAMLSS in R*, University of Lancaster, Lancaster, UK, 2009.
- [51] R. A. Rigby, *Distributions for Modeling Location, Scale, and Shape: Using GAMLSS in R*, Chapman and Hall/CRC, Boca Raton, FL, USA, 2019.

## Research Article

# Is It Possible to Earn Abnormal Return in an Inefficient Market? An Approach Based on Machine Learning in Stock Trading

Bui Thanh Khoa <sup>1</sup> and Tran Trong Huynh <sup>2</sup>

<sup>1</sup>Finance, Economics, and Management Research Group, Ho Chi Minh City Open University, Ho Chi Minh City, Vietnam

<sup>2</sup>FPT University, Hanoi, Vietnam

Correspondence should be addressed to Bui Thanh Khoa; khoa.bt@ou.edu.vn

Received 3 November 2021; Revised 22 November 2021; Accepted 25 November 2021; Published 8 December 2021

Academic Editor: Daqing Gong

Copyright © 2021 Bui Thanh Khoa and Tran Trong Huynh. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Risk management and stock investment decision-making is an essential topic for investors and fund managers, especially in the context of the COVID-19 pandemic. The problem becomes easier if the market is efficient, where stock prices fully reflect potential risk. Nevertheless, if the market is not efficient, investors may have an opportunity to find an effective investment method. Vietnam is one of the emerging markets; the efficiency is still weak. Thus, there will be an opportunity for astute investors. This study aims to test the weak-form efficient market and provide a modern approach to investors' decision-making. To achieve that aim, this study uses historical data of stocks in the VN-Index and VN30 portfolio to buy and sell within a one-day period under the rolling window approach to test the Ho Chi Minh City Stock Exchange (HoSE) through a runs test and to perform stock trading using the support vector machine (SVM) and logistic regression. The buying/selling of stocks is guided by the forecasted outcomes (increase/decrease) of logistic regression and SVM. This study adjusted the return rate in proportion to the risks and compared it with index investments of VN-Index and VN30 to evaluate investment efficiency. The test results dismissed the weak-form efficient-market hypothesis, which opens up many opportunities for short-term traders. This study's primary contribution is to provide a stock trading strategy for short-term investors to maximize trading profits. Because logistic regression and SVM have proven effective trading methods, investors can use them to achieve abnormal returns.

## 1. Introduction

Risk management and stock investing decision-making are critical topics for investors and fund managers, particularly regarding the COVID-19 pandemic. If the market is efficient, where stock prices adequately represent a possible risk, the issue becomes simpler to solve [1, 2]. Some investors often use technical analysis to select stocks as historical data (mainly price and trading volume) in the short term. Some technical analysis tools forecast price movement direction, deciding whether to buy or sell stocks [3]. Mizrach and Weerts [4] used technical indicators, price, and volume history to forecast future stock returns, sometimes called "chartists" because they use graphical trading representations. Azzopardi [5] applied principles to study how human emotions impact financial decision-making. SVM and

artificial neural networks (ANN) identify market abnormalities in many financial markets worldwide [6]. Nevertheless, Fama [7] proposed efficient-market hypothesis casts doubt on the reliability of the technical analysis. This theory will not help beat the market because it assumes that the price of a security fully reflects all available information [8–10]. That said, each market is efficient to a certain extent; specifically, there are three types of efficient markets in ascending order: weak, semistrong, and strong. Even in the weak form, the stock's price fully reflects its historical data.

For that reason, the security price cannot be predicted solely based on past prices [11]. Some empirical evidence suggests that markets are not truly efficient, which implies that investors may use templates or prediction models to achieve a higher rate of return [12, 13]. Hawaldar et al. [14] tested the weak-form efficient-market hypothesis of the



Bahrain Bourse stock market for the period 2011 to 2015 and concluded that the Kolmogorov–Smirnov goodness-of-fit test, run test, and autocorrelation test reject the weak-form efficient-market hypothesis. Kumar et al. [15] supported India’s weak-form efficient-market hypothesis for 2012–2017 but rejected the medium-form efficient-market hypothesis. Mensi et al. [16] studied the daily closing prices on the global and regional GIPSI stock markets in the USA and five GIPSI stock markets in Europe from January 1, 2009, to September 8, 2017. GIPSI, worldwide, and US markets are all inefficient, particularly in the short term. Whatever the time range, the Greek stock market is the most inefficient of all markets. In the short and long run, Portugal and Ireland have the least inefficient marketplaces. These findings also suggest that stock markets may not be suitable for risk diversification in asset allocation or risk hedging. The author also suggests that these findings have significant consequences for investors and policymakers. In reality, investors may utilize knowledge about long-term memory and the differential threshold for persistence across time horizons to outperform the market and generate abnormal returns.

A recent trend in behavioral finance theory is to explain that anomalies complement the shortcomings of the efficient-market hypothesis. Kahneman and Tversky, a pioneering researcher, point out that investors rely heavily on emotions and instincts rather than rationality to make decisions [17]. Emotional decision-making can lead to mistakes when making irrational investment choices. Some anomalies associated with behavioral finance theory include calendar, fundamental, and technical anomalies [18]. Some experiments show the weekend effect, holiday effect, turn-of-the-month effect, and January effect [19]. Rossi studied calendar anomalies in the Milan Stock Exchange from January 2005 to December 2015. They found that returns were negative on Monday and positive on Wednesday. Thus, investors should buy on Monday and sell on Wednesday. One limitation of these studies is that the effects may disappear or even reverse [20]. As a result, investors may be exposed to risks when using these investment trends.

This study aims to test the weak-form efficiency of the HoSE market and determine whether investors using logistic regression and the SVM model can outperform the market. The runs test approach rejects the weak form of an efficient market. These findings suggest that classic econometric and statistical models are likely to beat the market. However, the constantly evolving machine learning algorithms provide a viable alternative to traditional regression models. Some studies on the SVM application in finance have obtained many positive results, such as Cao and Tay [21], Huang et al. [22], Lu et al. [23], Mohamed [24], Azimi-Pour et al. [25], and Syriopoulos et al. [26]. The rolling window drives the buying and selling of securities by the logistic regression model’s output and the SVM algorithm. Input variables include close (closing price); HL (the highest minus lowest price); LO (lowest price minus opening price); variation (the difference in closing price between 2 consecutive trading sessions); ma7, ma14, and ma21 (average price of 7, 14, and 21 consecutive sessions, respectively); sd7 (standard

deviation of 7 consecutive sessions); vnc (the difference in closing prices of VN-Index for 2 consecutive sessions); vnipc (return rate of VN-Index portfolio); and insect (time trend). The data covers all stocks in the VN30 basket from January 28, 2000, to July 30, 2021. As a result, the SVM investment strategy beat the market with an extremely high average return rate.

Machine learning may discover weak-form efficient markets and develop trading methods for short-term investors, thereby maximizing earnings. Predicting the movement of stock prices using algorithms, such as the SVM model, has demonstrated a high accuracy. The parameters of the machine learning model were accurately predicted using the rolling window technique. Since a sample’s representativeness may be impaired by a period too short or too lengthy, 365 days is a good choice for a historical data set. Stock investing in a weak market is usually tricky for short-term investors. The SVM model, in particular, is a valuable tool for predicting the direction of price movement in the market. It is necessary to modify the investment returns to reflect the inherent risks to raise the degree of trust in the investment performance review. The Sharpe ratio is used to manage risk, while the *T*-test is used to evaluate trading methods. Due to the SVM model’s superb accuracy, the trading strategy employing it has produced a great return.

The following diagram depicts the flow of this study. Next, a brief review of relevant literature is provided: the efficient-market hypothesis (EMH), logistic regression, support vector machine (SVM). Section 3 of the study provides the conceptual foundation for the paper, including the theories of weak-form efficient-market hypothesis testing and price movement forecasting decision. Section 4 focuses on empirical data and outcomes. Section 5 provides further in-depth explanations of the study’s findings. In Section 6, the conclusions of this study and the limits and potential for further research are summarized and explained.

## 2. Literature Review

*2.1. Efficient-Market Hypothesis (EMH).* Fama [11] first proposed EMH in the 1970s. This article is significant because it paved the way for many other studies on the accuracy of the EMH theory. The concept of efficiency refers to the rapid absorption of information instead of the resources that produce maximum output as in other fields of economics. Information is defined as news that can affect prices and is unpredictable. In capital markets, efficient markets can be interpreted in various ways. The market in which prices always reflect available information is called an efficient market [11]. Meanwhile, Malkiel [27] argued that a capital market is efficient if it wholly and correctly reflects all relevant information in determining security prices. Generally, however, markets are considered efficient for certain types of information if disclosing that information to participants does not affect stock prices. EMH includes the following hypotheses:

Weak-form efficiency hypothesis: this degree of efficiency exists when a security’s price reflects historical data



about a security's price, including stock price and trading volume. In other words, one can forecast current stock prices on past stock prices. Testing the weak-form efficient-market hypothesis mainly concerns whether there is a statistical dependence between price changes. In other words, if the price changes are random, the market is a weak-form efficient market. Several frequently used testing techniques are autocorrelation and Ljung–Box's Q [28], variance ratio, LM-test [29], CD-test [30], Wright's test [31], runs test, January effect, and unit root test [32].

Semistrong-form efficiency hypothesis: this degree of efficiency exists when a security's price reflects publicly accessible market information, including historical data on security prices and publicly available information in the market, such as those in an issuer's prospectus. The semistrong-form efficient market encompasses the weak-form hypothesis because all market information, including stock prices, interest rates, and trading volume, must be publicly analyzed using the weak-form efficient-market hypothesis. Public information includes all nonmarket data, such as earnings and dividend announcements, P/E ratio, D/P ratio, P/B ratio, stock splits, and political economy information. Studies examining semistrong-form EMH can be classified into these two categories:

- (i) Studies that sought to forecast future rates of return using publicly accessible data, except for pure market data such as price levels and trading volumes, have been included in the weak-form test. These studies may include time series analysis of returns or cross-sectional distribution of returns of individual stocks. EMH proponents argue that it is impossible to use publicly available information to predict future returns using past returns or to forecast future cross-sectional distributions of returns (e.g., highest quartiles or deciles of returns) [33–36].
- (ii) Event studies investigate how quickly stock prices change in response to specific key economic events. One practical approach is to test whether it is feasible to invest in stocks and earn an extraordinarily high rate of return after a significant event (such as stock merges, stock splits, central economic data, and principal) is publicly announced or not. Again, EMH proponents expect stock prices to adjust rapidly so that investors cannot earn high returns by buying after public announcements and paying regular transaction costs [37–40].

Strong-form efficiency hypothesis: this degree of efficiency exists as all information is fully reflected in stock prices, including nonpublic information such as internal information. The strong-form efficient-market hypothesis combines both the weak-form and the semistrong-form efficient hypothesis. The strong-form efficient-market hypothesis extends the assumption of efficient markets, in which prices reflect publicly available information to a perfect market, and all information is free and available. It is necessary to know when internal or insider information arises to evaluate strong-form efficient markets. This timing

is hard to identify. Strong-form efficient markets are often researched in developed countries. For emerging markets, most studies focus on weak- and semistrong-form EMH. The exploration of strong form effectiveness is still a controversial matter among scholars [41–43].

**2.2. Logistic Regression.** Logistic regression is a statistical technique that describes the relationship between independent variables and binary dependent variables (which can also be applied to discrete dependent variables). Through this relationship, logistic regression allows the output prediction of a given set of input values. In predicting the output using logistic regression, this study calculates the probability that the output takes the value 1 with the given observation data to find  $P(Y = 1|X)$ . With the assumption of binomial distribution of the dependent variable, this study considers the odd ratio as follows:

$$G(X) = \frac{P(Y = 1 | X)}{P(Y = 0 | X)} = \frac{P(Y = 1 | X)}{1 - P(Y = 1 | X)}. \quad (1)$$

Taking the logarithm on both sides of (1), this study has

$$\ln G(X) = \ln \left( \frac{P(Y = 1 | X)}{1 - P(Y = 1 | X)} \right) = X\beta, \quad (2)$$

where  $\beta = (\beta_0, \beta_1, \dots, \beta_k)$  that are the parameters to be estimated.

From equation (2), this study makes the equivalent transformation as follows:

$$P(Y = 1 | X) = \frac{e^{X\beta}}{1 + e^{X\beta}}. \quad (3)$$

Usually, the maximum likelihood estimation (MLE) method is used to estimate the parameter  $\beta$ . The classification rule is determined by equation (3) as follows:

$$y_i = \begin{cases} 1, & P(y_i = 1 | X) \geq 0.5, \\ 0, & P(y_i = 1 | X) < 0.5. \end{cases}$$

Logistic regression is applied in many fields for the binary dependent variable. In finance, Han et al. [44] used a sample of 76 firms and 32 variables related to their financial ratios to predict precarious financial situations. The authors used the backward stepwise method in logistic regression and obtained results with high accuracy of 92.86%. Konglai and Jingjing [45] used logistic regression to analyze listed companies' credit risk in China. The data used included 130 companies with 6 dependent variables and was divided into 90 companies for the training set and 40 for the testing set. The training sample has an accuracy of 87.8%, while the testing set has a precision of 75%. Table 1 summarizes some publications that have used typical logistic regression.

**2.3. Support Vector Machine (SVM).** The SVM algorithm was proposed by Vapnik and Lerner [50] to solve the classification issue. SVM is a supervised mathematical algorithm used to classify data in different dimensions. Suppose that  $Y$

TABLE 1: Logistic regression in prior studies.

Authors	Research	Accuracy
Jabeur [46]	Bankruptcy prediction using partial least squares logistic regression	~70%
Rafatnia et al. [47]	Financial distress prediction across firms	83.3%
Jovanović et al. [48]	Financial indicators as predictors of illiquidity	95.5%
Strzelecka et al. [49]	Application of logistic regression models to assess household financial decisions regarding debt	70.5%

is a categorical variable with two possible values  $-1$  and  $1$  and  $X$  is an input variable. The classification hyperplane is defined by the equation:  $wx^T + b = 0$ , where  $w$  and  $b$  are the coefficients. The coefficients  $w$  and  $b$  should be chosen such that  $wx^T + b \geq 1$  if  $y_i = 1$  and  $wx^T + b \leq -1$  if  $y_i = -1$ . The training set is used to find  $w$  and  $b$  such that  $\|w\|$  is minimized, and the vectors  $x_i$  in which  $|y_i|(wx_i^T + b) = 1$  are called support vectors. To improve classifier efficiency, a kernel function is used to map the data to a high-dimensional space where the data will be more clearly segregated. The kernel function is defined by the dot product:  $K(x, y) = \langle f(x), f(y) \rangle$ . Some common kernel functions are linear, polynomial, and radial basic function. Nevertheless, for some complex data sets, it is impossible to find a perfect hyperplane. Hence, Cortes and Vapnik [51] propose to add soft margins, that is, accepting some misclassified observations. The SVM algorithm is now minimized:  $\min_{w,b,\xi} (1/2w^T w + C \sum_{i=1}^n \xi_i)$  given that  $y_i(w^T w\phi(x_i) + b) \geq 1 - \xi_i$ , where  $C$  is a hyperparameter and  $\phi$  is a conversion mapping from low- to high-dimensional space.

SVM is often used in financial research. For instance, Kim [52] has used SVM to predict hotels' bankruptcy in Korea. Between 1995 and 2002, a sample of 33 hotels was collected, and the forecast results achieved 95% accuracy. In the Japanese market, Huang et al. [22] used SVM to predict the direction of the NIKKEI 225 Index and showed that SVM outperformed other classification methods in their study, including random walk model, quadratic discriminant analysis (QDA), and ANN. Ren et al. [45] integrated SVM with investor behavior analysis in the Chinese market. This study forecasted the SSE 50 Index's movement from 2014 to 2016 in 485 trading days, used both fivefold and rolling window methods, and reached a maximum accuracy of 89.93%.

### 3. Research Data and Methods

**3.1. Research Data and Variable Description.** Research data includes 30 companies in the VN30 basket (unadjusted price), VN-Index, and VN30 index in a one-day period. Table 2 describes the tickers and their observations in the VN30 basket.

The data collection period was from July 28, 2000, to July 30, 2021, in which some companies were newly established, and there were some days off. Hence, the number of observations of these companies was varied. The data was collected from the website <https://www.hsx.vn> (Ho Chi Minh City Stock Exchange). Each observation included date, ticker, closing price, opening price, highest price, lowest price, and trading volume. The variables in the study are described in Table 3.

TABLE 2: Tickers and observations in the VN30 basket.

Ticker	Observations	Ticker	Observations	Ticker	Observations
BID	1,840	MWG	1,731	TCB	762
BVH	2,989	NVL	1,112	TCH	1,173
CTG	2,975	PDR	2,573	TPB	789
FPT	3,611	PLX	1,039	VCB	2,987
GAS	2,265	PNJ	3,043	VHM	772
HDB	857	POW	605	VIC	3,424
HPG	3,381	REE	5,050	VJC	1,076
KDH	2,833	SBT	3,319	VNM	3,838
MBB	2,401	SSI	3,600	VPB	957
MSN	2,896	STB	3,720	VRE	898

### 3.2. Research Method

#### 3.2.1. Testing the Weak-Form Efficient-Market Hypothesis.

According to the weak-form efficient-market theory, a security's past prices cannot forecast current prices and generate abnormal returns. There are other testing techniques available, but these studies employ runs tests like some previous studies, including Fawson et al. [57], Moustafa [58], Ahmad et al. [59], Nisar and Hanif [60], Hamid et al. [61], and Wei [62]. The runs test, known as the Wald-Wolfowitz test, is a nonparametric statistical test that examines the randomness hypothesis on a two-state data series [63]. The runs test will assess whether the elements of the series appear independently. In other words, if assuming the price increases or stays the same as (+) and decrease as (-), then a weak-form efficient market implies that price changes are independent. When the sample size is large enough, the statistic  $Z = R - \bar{R}/s_R \sim N(0, 1)$ , where:

$R$ : number of runs in the sample (each run is a sequence of consecutive "+" or "-" signs)

$\bar{R}$ : expected value of  $R$ , calculated by the formula  $\bar{R} = 2n_1n_2/n_1 + n_2 + 1$

$s_R^2$ : the standard error of the runs,  $s_R^2 = 2n_1n_2(2n_1n_2 - n_1 - n_2)/(n_1 + n_2)^2(n_1 + n_2 - 1)$  ( $n_1, n_2$  are the number of "+" and "-" signs, respectively)

Mainly this method explores the randomness in the changes of the VN-Index and VN30 index. If this variation is random, it supports the weak-form efficient-market hypothesis, suggesting that traditional forecasting models using historical data are unlikely to produce an excess return.

Finally, to test the influence of factors affecting price movements, we performed logistic regression for all data in the research period. This result also implies that investors with little experience in academic knowledge can still base

TABLE 3: Variable description.

Variables	Formula	Description	Source
$close_t$		The closing price at date $t$	Schöneburg [53]
$fore\ dir_t$	$fore\ dir_t = \begin{cases} 1, & close_t \geq close_{t-1} \\ -1, & close_t < close_{t-1} \end{cases}$	The direction of price movement (foredir = 1 implies that the closing price increases from the previous day)	Ren et al. [54]
$HL_t$	$high_t - low_t$	The fluctuation of the price within a trading day	
$LO_t$	$low_{t-1} - open_t$	The difference between the lowest price and the opening price	
$variation_t$	$close_t - close_{t-1}$	The fluctuation of closing prices between 2 consecutive days	
$ma7_t$	$1/7 \sum_{i=0}^6 close_{t-i}$	The average closing price of 7 consecutive trading sessions	Vijh et al. [55]
$ma14_t$	$1/14 \sum_{i=0}^{13} close_{t-i}$	The average closing price of 14 consecutive trading sessions	
$ma21_t$	$1/21 \sum_{i=0}^{20} close_{t-i}$	The average closing price of 21 consecutive trading sessions	
$sd7_t$	$\sqrt{\text{var}(close_t, close_{t-1}, \dots, close_{t-6})}$	The standard deviation of the closing price of 7 consecutive trading sessions	
$vnic_t$	$vmin\ de\ x_t - vmin\ de\ x_{t-1}$	Fluctuation of VN-Index between 2 consecutive trading sessions	Qiu and song [56]
$vnipc_t$	$vmin\ de\ x_t - vmin\ de\ x_{t-1} / vmin\ de\ x_{t-1} \times 100$	The return rate of the VN-Index portfolio	
$insec_t$		Time trend variable (the default origin is January 1, 1970)	

the fluctuations of variables (variables with strong impact) to make investment decisions.

**3.2.2. Price Movement Forecasting and Investment Decision-Making.** This study focuses on two models, logistic regression and SVM, to forecast price movement direction. Assuming that the historical data has a maximum value of 1

year, the study will use fixed training data of 365 observations to make forecasts using the “rolling window” method. Algorithms are used to identify the optimal parameters for the first 365 observations, forecast the 366th observation, and continue until the last observation, as shown in Figure 1.

### 3.3. Forecasting Model

$$\widehat{foredir}_{t+1} = f(close_t, HL_t, LO_t, variation_t, ma7_t, ma14_t, ma21_t, sd7_t, vnic_t, vnipc_t, insec_t). \quad (4)$$

The sigmoid function is employed for the logistic regression model, and the MLE method is used to estimate the regression coefficients. For the SVM algorithm, the kernel function radial and  $\gamma = 0.1$  are used. Based on the logistic regression and SVM models, the investors will buy/sell stocks, respectively. To assess investment performance, this study adjusts risks using the Sharpe ratio [64, 65]:  $\text{SharpeRatio} = r_p - r_f / \sigma_p$ , where

$r_p$ : return rate of the portfolio (or security)  $p$

$r_f$ : risk-free rate (1-year treasury note)

$\sigma_p$ : standard deviation of the portfolio (or security)  $p$

Finally, this study compares the performances of investments made by the logistic regression model and SVM with investments made by the  $T$ -test according to VN30 and VN-Index. Furthermore, this study seeks to determine whether holding a single stock is more efficient than holding a market portfolio index. The novelty of this study is to provide a securities trading method using a logistic regression model and SVM.

## 4. Results

**4.1. Descriptive Statistics.** The descriptive statistics of the variables are described in Table 4 below. The table shows that the price fluctuates from 0.233 USD/share to 23.233 USD/share; the average price is 2.266 USD/share. The foredir ends up with 39,096 observations resulting in a decrease in closing price compared to the day before. The remaining 29,420 observations of closing price were not decreased; the specific amount is shown in Figure 2. The most muscular daily closing price movement-down 7.108 USD/share on a day, occurred to VNM ticker on July 5, 2007 (exchange rate USD/VND = 22,748).

The fluctuations of the variables close and variation are better shown in the boxplot on Figures 3 and 4. Some tickers such as FPT, REE, SSI, and STB tilted to the right and had unusually high closing prices in some trading sessions. Still, the tickers' variation is mostly stable. This study noticed an anomaly that FPT plummeted 7.429 USD/share on May 21, 2007, the most profound fall across all stocks in the VN30 portfolio in all trading sessions. The decline in share price is

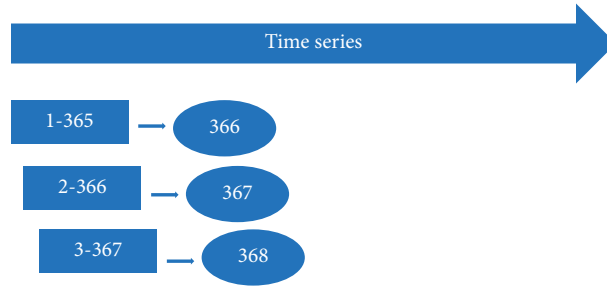


FIGURE 1: Rolling window method.

TABLE 4: Descriptive statistics of variables.

Statistics	Close	HL	LO	Variation	ma7	ma14	ma21	sd7
Min	5.3	-161.7	-40	-169	5.39	5.59	5.83	0
Median	37.2	0.8	-0.3	0	37.2	37.25	37.26	0.66
Max	665	44.5	82	30	634.57	623	622.14	93.83
Mean	51.54	1.242	-0.62	0	51.56	51.58	51.59	1.19
SD	43	1.6	1.13	2	43.1	43.2	43.27	2.03

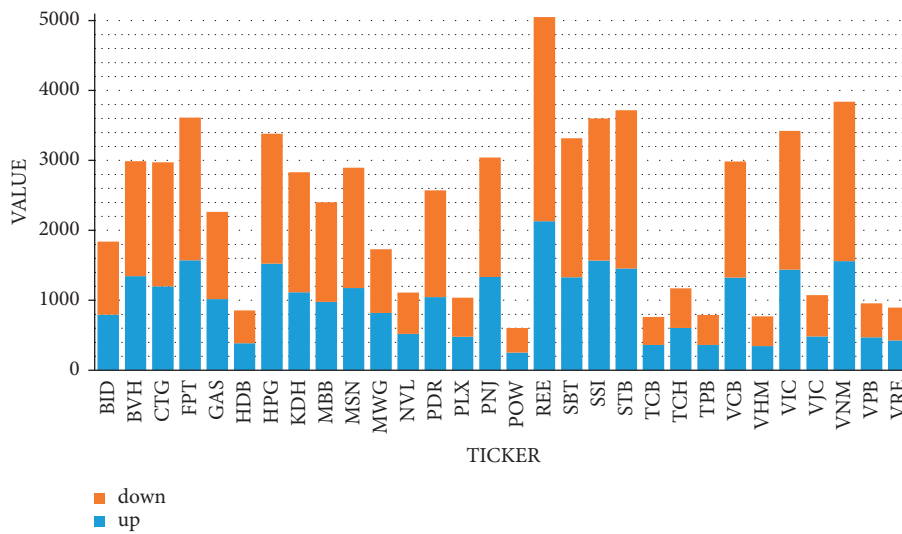


FIGURE 2: The number of observations on the tickers' price increases/decreases.

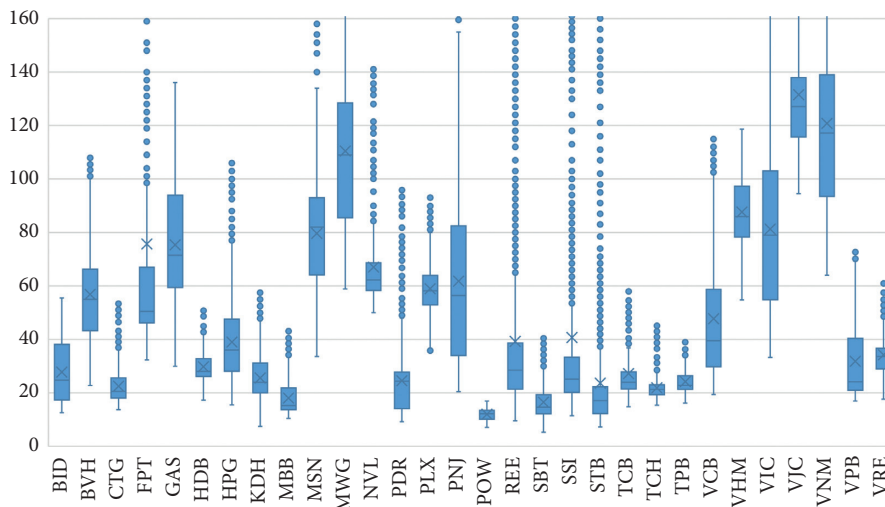


FIGURE 3: Closing prices movement of tickers.



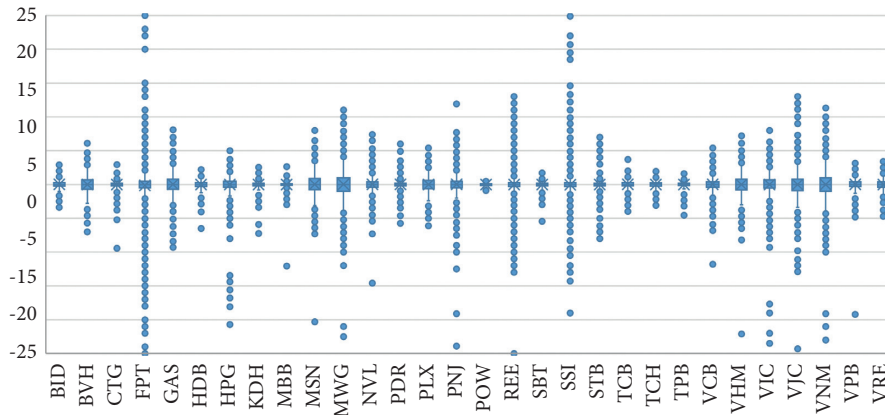


FIGURE 4: Variation movement of the tickers.

due to FPT's dividend payment policy with a payout ratio of 2:1, which shows that one more FPT share will be awarded for every two FPT shares an investor holds.

For the market, this study has a summary table detailing the variables closevn (closing price of VN-Index), vnic, vnipc, closevn30 (closing price of VN30 index), rvn30 (return rate of VN30 index), and rf (the interest of 1-year government bonds). Table 5 and Figure 5 show that the closing prices of the VN-Index and VN30 index primarily fluctuate together, while bond interests are primarily stable and tend to decrease. From the beginning of 2020, this study noticed that both the VN-Index and VN30 dropped significantly and then rose again. This result was because of the COVID-19 pandemic, which obstructed the production and trading activities of businesses. When the businesses stabilized, the cash flow poured into the financial investments, leading to increased stock prices.

**4.2. Runs Test Results.** Runs test results showed that the weak-form efficient-market hypothesis is dismissed at 1%, implying that technical analysis can obtain an abnormal return.

**4.3. Accuracy of Price Movement Forecasting Models.** This study used the logistic regression model and SVM to forecast the increase and decrease of stocks based on the rolling window method. The accuracy value (the number of correct predictions out of the total predictions) is summarized in Table 6. The average accuracy in forecasting 30 stocks of the logistic regression model and SVM are 58.93% and 92.48%. The SVM model has proven to be more effective than the logistic regression model.

**4.4. Stock Trading Results.** Stocks were traded on the stock price increase and decrease forecasts made by the logistic regression and the SVM models. The results of average daily return before and after risk adjustment are in Table 7. As seen in Tables 5 and 7, the SVM model outperforms the logistic regression model and the portfolio index investment (including VN30 and VN-Index). To determine the efficacy of the trading strategies, this study conducted five one-sided

*T*-tests with the null hypothesis (investments are not more efficient than index portfolio investments) and the alternative hypothesis (investment methods are more efficient). Table 8 summarizes the results of the tests by *p*-value. The terminologies in Tables 7 and 8 are explained in Table 9.

**4.5. Factor Affecting the Stock Price Movement.** This study performed logistic regression for the entire data to determine the factors affecting the stock price movement. Logistic regression results are shown in Table 10. The regression result in Table 10 shows that the factors HL, LO, variation, vnic, vnipc, insec, and sd7 have a statistically significant impact, of which vnipc has the most substantial impact. This conclusion shows that market portfolio return is the strongest indicator of price change expectations; for every extra percentage rise in market portfolio return, investors anticipate the odds ratio increasing by 0.2. In addition, the model also shows that the moving average indicators (MA) are not statistically significant at 0.1, that is, the MA indicator does not affect stock trading. Volatility indicators HL and LO have regression coefficients of 0.055 and 0.061, respectively. Both are statistically significant, showing that these fluctuations increase the possibility of bullish forecasting for the next trading session. Nevertheless, the vnic indicator has a negative coefficient and is statistically significant, showing that the greater the market volatility, the more it predicts that the price will decrease.

## 5. Discussion

The nonparametric runs test examines the randomness of a sequence of rising/falling states of stock prices. The weak-form efficient market implies that prices rise/fall randomly [66]. This study performs a runs test on two rising/falling ranges of the VN30 and VN-Index portfolios with the null hypothesis that the direction of price movement is random. Runs test results in Table 11 have a *p*-value less than 0.01. This study rejects the null hypothesis for both tests [67]. This result implies that the weak-form efficient-market hypothesis is rejected. This result is also consistent with some previous research [61, 68–70]. Market weakness is not guaranteed to present an opportunity for short-term traders

TABLE 5: The statistics of the variables in the market.

Statistics	closevn	vnic	vnipc	closevn30	rvn30	rf
Min	235.5	-73.23	-7.15	230.7	-6.73	0.27
Median	582	0.68	0.1	614.1	0.11	4.86
Max	1,420.3	40.85	4.74	1,557.4	5.16	13.83
Mean	668.7	0.32	0.04	677	0.06	5.77
SD	246.45	8.7	1.28	228.59	1.31	3.49

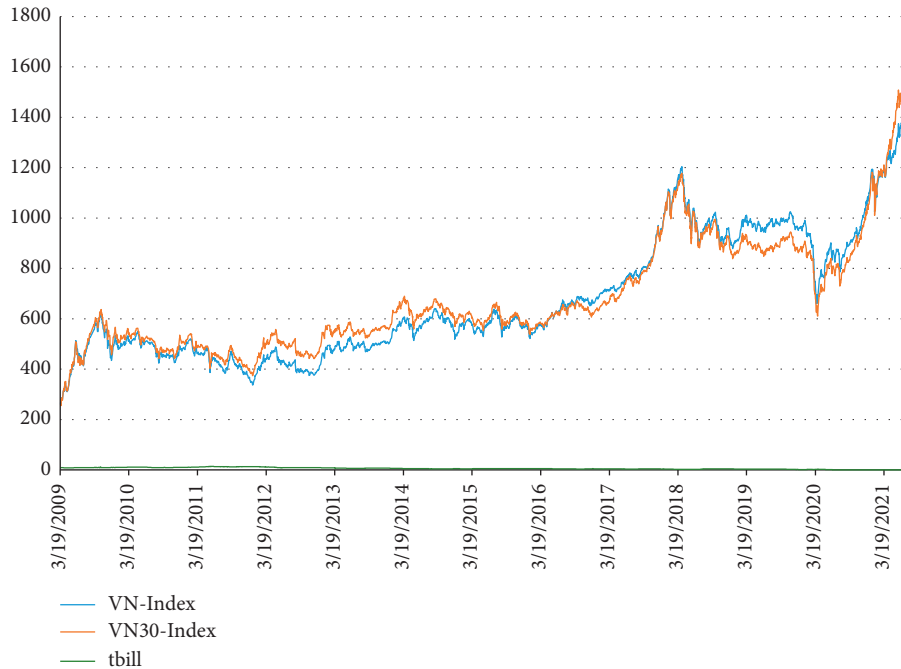


FIGURE 5: Movement of VN-Index, VN30 index, and bond interests (T-bill).

TABLE 6: A summary on accuracy of tickers in VN30 portfolio.

Ticker	Logistic	SVM	Ticker	Logistic	SVM	Ticker	Logistic	SVM
BID	59.49	92.55	POW	56.85	93.36	MWG	53.84	91.51
BVH	58.4	92.91	REE	62.21	92.53	NVL	58.29	93.72
CTG	61.93	93.53	SBT	62.2	93.6	PDR	59.98	92.76
FPT	58.67	90.64	SSI	58.9	91.75	PLX	56.3	91.41
GAS	57.55	92.9	STB	63.32	92.19	PNJ	57.86	91.64
HDB	57.61	93.71	TCB	55.53	95.73	VIC	59.87	91.57
HPG	58.93	91.22	TCH	54.88	91.97	VJC	58.57	86.66
KDH	61.89	93.36	TPB	58.82	96.94	VNM	61.83	91.34
MBB	62.15	93.67	VCB	58.6	93.25	VPB	57.5	88.03
MSN	62.48	93.05	VHM	57.84	94.85	VRE	55.62	91.95
Average accuracy			SVM	92.48	Logistic			58.93

looking for past patterns to rely on when buying/selling to maximize trading profits.

This study implements three trading strategies: the logistic regression model, the SVM model, and holding stocks for the long term. In the first two strategies, the models forecast the increase/decrease of the stock price, resulting in buying and selling correspondingly. Compared to the traditional logistic regression model, the SVM model better

predicts price movement direction. On all 30 tickers in Table 6, the SVM model defeated the logistic regression model. Additionally, its accuracy is exceptional, averaging 92.48% and 58.93%. This finding is much like prior studies, which show that SVM produces greater accuracy than the logistic regression model [71–74].

The accuracy of the SVM model in Table 6 is very high, with most of them correct over 90%, except for the two



TABLE 7: Stock trading results.

Ticker	SVM	Logistic	Return	Adj SVM	Adj logistic	Adj return
BID	1.448	0.177	0.067	0.780	0.072	0.025
BVH	1.450	0.280	0.038	0.800	0.116	0.010
CTG	1.432	0.380	0.046	0.788	0.162	0.014
FPT	1.353	0.331	0.054	0.726	0.135	0.015
GAS	1.421	0.257	0.068	0.775	0.108	0.026
HDB	1.469	0.465	0.059	0.839	0.207	0.025
HPG	1.333	0.316	0.048	0.705	0.132	0.014
KDH	1.427	0.427	0.039	0.781	0.183	0.012
MBB	1.462	0.398	0.067	0.811	0.170	0.025
MSN	1.434	0.449	0.038	0.793	0.194	0.011
MWG	1.359	0.211	0.094	0.716	0.088	0.037
NVL	1.411	0.284	0.083	0.808	0.125	0.035
PDR	1.494	0.396	0.062	0.812	0.166	0.022
PLX	1.240	0.226	0.065	0.702	0.104	0.028
PNJ	1.375	0.329	0.038	0.720	0.136	0.010
POW	1.344	0.317	0.034	0.888	0.158	0.017
REE	1.552	0.467	0.053	0.730	0.134	0.013
SBT	1.427	0.411	0.049	0.778	0.174	0.015
SSI	1.372	0.309	0.053	0.748	0.127	0.015
STB	1.402	0.425	0.052	0.721	0.160	0.014
TCB	1.603	0.271	0.006	0.864	0.111	0.001
TCH	1.475	0.242	0.056	0.784	0.100	0.021
TPB	1.643	0.419	0.027	0.936	0.176	0.010
VCB	1.452	0.325	0.041	0.803	0.136	0.012
VHM	1.590	0.477	0.035	0.859	0.198	0.014
VIC	1.396	0.369	0.049	0.758	0.156	0.015
VJC	1.268	0.367	0.047	0.690	0.165	0.019
VNM	1.437	0.460	0.063	0.743	0.176	0.014
VPB	1.316	0.305	0.045	0.755	0.139	0.019
VRE	1.397	0.355	0.071	0.802	0.160	0.030
Average	1.426	0.348	0.052	0.781	0.146	0.018

tickers: the VJC ticker and VPB ticker. Moreover, its lowest accuracy is 86.66%, and the highest is 96.94% (for TPB ticker). This result is better than similar studies such as Kim [75], Kara et al. [76], Patel et al. [77], and Duong et al. [78]. One success in the SVM model comes from its model estimation method. Compared to other methods, the “rolling window” is more efficient because the continuous-time series ensures the input parameters’ accuracy. The 365-day period is a reasonable choice. If it is longer, the data will become too outdated. If it is shorter, the collected data may not be a good representation of the whole. Specifically, the training data is permanently fixed for the latest 365 observations. Because of the continual updating of the training data set, the initial parameters are adjusted accordingly, increasing the forecasts’ accuracy.

In contrast, the sample’s representativeness will be a problem if the data set is split into two independent sets. For example, Vijn et al. [55] divided the data set into two sets: the training data set (June 4, 2009–March 4, 2017) and the testing data set (April 4, 2017–May 4, 2019). The parameters calculated by the training data set are too outdated for forecasting; using data from 2017 to forecast for 2019 does not seem to be reasonable. Cao and Tay [21] and Ji et al. [79] divided the data set into three sets: training, validation, and testing data. While rationality is much better when applied

historical data, performance will be significantly less than the rolling window.

The superior predictive power of the SVM model has led to excellent trading performance. From Table 7, using the SVM model for trading has achieved an average rate of return of 1.426%/day with the corresponding Sharpe ratio of 0.781, which is much greater than using the logistic regression model. Although the logistic regression method is not as effective as the SVM model, it still produces a great result with an average return rate of 0.348%/day and a Sharpe ratio of 0.146. In contrast, the average rate of return of VN30 and VN-Index is only 0.06% per day and 0.04% per day, respectively. The efficiency test results of all three methods (trading under the SVM model, logistic regression, and long-term holding of individual stocks) in Table 8 suggest that the SVM method is more efficient than investment according to the VN30 and VN-Index with a significance level of 0.001 (the  $p$ -values are approximately 0). Trading using the logistic regression model is effective when 25 out of 30 stocks achieved statistical significance at 0.1. For long-term holding of individual stocks, the average return rate is 0.052%/day, higher than VN-Index (0.04%/day) but lower than the VN30 index (0.06%/day). Furthermore, the  $p$ -values are all greater than 0.1, implying that the investing strategy of long-term holding individual stocks cannot outperform the market.

TABLE 8: The  $p$ -value for testing the effectiveness of trading methods.

Ticker	p_LR_vn30	p_LR_vni	p_SVM_vn30	p_SVM_vni	p_return
BID	0.291	0.178	0.001	0.001	0.641
BVH	0.001	0.001	0.001	0.001	0.549
CTG	0.001	0.001	0.001	0.001	0.499
FPT	0.001	0.001	0.001	0.001	0.543
GAS	0.020	0.010	0.001	0.001	0.578
HDB	0.020	0.004	0.001	0.001	0.574
HPG	0.001	0.001	0.001	0.001	0.464
KDH	0.001	0.001	0.001	0.001	0.628
MBB	0.001	0.001	0.001	0.001	0.531
MSN	0.001	0.001	0.001	0.001	0.643
MWG	0.224	0.131	0.001	0.001	0.581
NVL	0.069	0.027	0.001	0.001	0.428
PDR	0.001	0.001	0.001	0.001	0.666
PLX	0.211	0.103	0.001	0.001	0.550
PNJ	0.001	0.001	0.001	0.001	0.599
POW	0.615	0.387	0.001	0.001	0.897
REE	0.001	0.001	0.001	0.001	0.562
SBT	0.001	0.001	0.001	0.001	0.346
SSI	0.001	0.001	0.001	0.001	0.542
STB	0.001	0.001	0.001	0.001	0.549
TCB	0.377	0.178	0.001	0.001	0.730
TCH	0.092	0.038	0.001	0.001	0.419
TPB	0.076	0.021	0.001	0.001	0.651
VCB	0.001	0.001	0.001	0.001	0.533
VHM	0.055	0.014	0.001	0.001	0.667
VIC	0.001	0.001	0.001	0.001	0.540
VJC	0.012	0.003	0.001	0.001	0.510
VNM	0.001	0.001	0.001	0.001	0.549
VPB	0.096	0.030	0.001	0.001	0.572
VRE	0.088	0.024	0.001	0.001	0.552

TABLE 9: Glossary of variables.

Variable	Definition
SVM	The average rate of return using the SVM model
Logistic	The average rate of return using the logistic regression model
Return	The average rate of return on investment for holding securities
Adj SVM, adj logistic, and adj return	Risk-adjusted rate of return
p_LR_vn30, p_LR_vni, p_SVM_vn30, and p_SVM_vni	The $p$ -value for testing if LR and SVM methods are more effective than investing by VN30 index and VN-Index
p_return	The $p$ -value for testing the efficiency of holding a single stock compared to holding the VN-Index

TABLE 10: Logistic regression result.

Min	Deviance residuals				Max
	1Q	Median	3Q		
-2.2376	-1.0705	-0.9763	1.2765	1.9416	
	Estimate	Std. error	z value	Pr(> z )	
(Intercept)	-1.54e + 00	8.77e - 02	-17.589	<2e - 16***	
close	3.57e - 03	4.41e - 03	0.81	0.4179	
HL	5.53e - 02	9.29e - 03	5.951	2.67e - 09***	
LO	6.07e - 02	1.11e - 02	5.47	4.50e - 08***	
variation	1.00e - 02	6.00e - 03	1.67	0.0949	
vnicbb	-2.51e - 02	2.18e - 03	-11.543	<2e - 16***	
vnipc	2.00e - 01	1.63e - 02	12.295	<2e - 16***	
insec	7.45e - 05	5.18e - 06	14.389	<2e - 16***	
ma7	4.88e - 03	7.81e - 03	0.625	0.5321	
ma14	-1.34e - 03	9.15e - 03	-0.146	0.8838	
ma21	-8.35e - 03	5.35e - 03	-1.562	0.1184	
sd7	3.98e - 02	5.54e - 03	7.184	6.76e - 13***	

Note. Significance codes: 0 " \* \* \* " 0.001 " \* \* " 0.01 " \* " 0.05 " . " 0.1 " " 1.

TABLE 11: Runs tests results.

VN30 index	VN-Index
Approximate runs test data: VN30	Approximate runs test data: VN-Index
Runs = 1,482; $p$ -value = 0.006332; alternative hypothesis: two-sided	Runs = 1,466; $p$ -value = 0.001233; alternative hypothesis: two-sided

Logistic regression results reveal that indicators such as HL, LO, variation, vnnc, vnipc, and sd7 impact stock price movement. Specifically, the increase of HL, LO, vnipc, and sd7 predicts that the price will increase, and VNC ticker increase predicts that the price will decrease. Indicators related to MA and close are not statistically significant and therefore do not have a predictive function of stock price movement.

## 6. Conclusion

Financial markets are efficient when old and new information is quickly reflected in the current price of a security. Therefore, because the current price includes historical information, technical analysis will not guarantee an excess return. Unfortunately, the test results reveal that the HoSE market is inefficient, meaning that technical analysis might generate abnormal returns.

The study's main contributions are identifying weak-form efficient markets and providing trading strategies for short-term investors by applying the machine learning model to optimize profits. Stock price movement forecasting algorithms, particularly the SVM model, have shown the predicting effectiveness, with an average accuracy of up to 92.48% and the peak accuracy of 96.94% (for the ticker TPB). The rolling window approach performed well in predicting the parameters of the machine learning model. The duration of the historical data is critical because the sample's representativeness may be compromised by a period that is too short or too long; hence, 365 days is considered a suitable option. Stock trading in an underperforming market is always a challenge for short-term investors. One trading strategy investors should consider there is the logistic regression model (especially the SVM model) to forecast price movement direction. Because high investment returns often conceal underlying risks, investment results should be adjusted accordingly to increase the confidence level in the investment performance evaluation. This study chooses the Sharpe ratio for risk adjustment and uses the  $T$ -Test to determine the effectiveness of trading strategies. Due to the high accuracy of the SVM model, the trading strategy using it has earned an exceptional rate of return.

Moreover, as the HoSE stock market is inefficient, short-term investors can rely on past patterns to maximize returns in trading. Short-term investors should consider using the SVM model and logistic regression models when making buying/selling decisions. The decision to choose trading stocks should be based on several indicators such as intraday price movement, price movement between two consecutive trading sessions, moving average, the standard deviation of the stock, and market volatility. It is possible to synthesize the SVM model from those indicators into an indicator for the final forecast. For long-term investors, it is better to

invest in a diversified portfolio or a portfolio index rather than holding individual stocks. Medium- to long-term investors should invest in a diversified portfolio or use fundamental analysis to select good stocks for a longer-term plan. Investors with limited knowledge related to pattern analysis can rely on indicators such as intraday price movement, price movement between two straight days, market volatility, and the stock's overall risk in the short term to forecast an increase or decrease in a security's price. Moreover, the return on the market portfolio is the most potent indicator because it reflects an optimistic attitude towards the market. If returns are positive, investors are more optimistic about market growth and thus decide to buy more; as a result, the stock price will increase.

Although this trading method has obtained an unprecedented return rate on short-term trading, the study omitted several factors such as transaction costs, taxes, and liquidity risk. Superior returns also use historical information, which is only valuable for inefficient market conditions. Therefore, more experiments are needed on inefficient markets to increase the reliability of the model. Further research may expand in two directions. First, the model's effectiveness in different markets has to be tested, and other factors such as tax transaction costs has to be considered. Second, the other authors can apply machine learning algorithms such as tree decision, deep learning, and neural networks to increase the model's predictive ability.

## Data Availability

The data are available on request.

## Conflicts of Interest

The authors declare that no known conflicting financial interests or personal connections have affected the work described in this article.

## Acknowledgments

This research did not receive any specific grant from funding agencies in the public or commercial sectors.

## References

- [1] R. Vanaga and B. Sloka, "Financial and capital market commission financing: aspects and challenges," *Journal of Logistics, Informatics and Service Science*, vol. 7, no. 1, pp. 17–30, 2020.
- [2] L. Zhang and H. Kim, "The influence of financial service characteristics on use intention through customer satisfaction with mobile fintech," *Journal of System and Management Sciences*, vol. 10, no. 2, pp. 82–94, 2020.

- [3] B. T. Khoa and D. T. Thai, "Capital structure and trade-off theory: evidence from vietnam," *The Journal of Asian Finance, Economics, and Business*, vol. 8, no. 1, pp. 45–52, 2021.
- [4] B. Mizrach and S. Weerts, "Highs and lows: a behavioural and technical analysis," *Applied Financial Economics*, vol. 19, no. 10, pp. 767–777, 2009.
- [5] P. V. Azzopardi, *Behavioural Technical Analysis*, Harriman House Limited, Petersfield, UK, 2010.
- [6] M.-W. Hsu, S. Lessmann, M.-C. Sung, T. Ma, and J. E. V. Johnson, "Bridging the divide in financial market forecasting: machine learners vs. financial economists," *Expert Systems with Applications*, vol. 61, pp. 215–234, 2016.
- [7] E. F. Fama, "The behavior of stock-market prices," *The Journal of Business*, vol. 38, no. 1, pp. 34–105, 1965.
- [8] S. S. Alexander, "Price movements in speculative markets: trends or random walks," *Industrial Management Review*, vol. 2, no. 2, p. 7, 1986.
- [9] E. F. Fama, "Random walks in stock market prices," *Financial Analysts Journal*, vol. 51, no. 1, pp. 75–80, 1965.
- [10] G. E. Metcalf and B. G. Malkiel, "The wall street journal contests: the experts, the darts, and the efficient market hypothesis," *Applied Financial Economics*, vol. 4, no. 5, pp. 371–374, 1994.
- [11] E. F. Fama, "Efficient capital markets: a review of theory and empirical work," *The Journal of Finance*, vol. 25, no. 2, pp. 383–417, 1970.
- [12] A. Shleifer, *Inefficient Markets: An Introduction to Behavioural Finance*, Oup Oxford, Oxford, UK, 2000.
- [13] R. Ball, J. Gerakos, J. T. Linnainmaa, and V. Nikolaev, "Earnings, retained earnings, and book-to-market in the cross section of expected returns," *Journal of Financial Economics*, vol. 135, no. 1, pp. 231–254, 2020.
- [14] I. T. Hawaldar, B. Rohith, and P. Pinto, "Testing of weak form of efficient market hypothesis: evidence from the Bahrain Bourse," *Investment Management and Financial Innovations*, vol. 14, no. 2, pp. 376–385, 2021.
- [15] A. Kumar, R. Soni, I. T. Hawaldar, M. Vyas, and V. Yadav, "The testing of efficient market hypotheses: a study of Indian pharmaceutical industry," *International Journal of Economics and Financial Issues*, vol. 10, no. 3, pp. 208–216, 2020.
- [16] W. Mensi, A. K. Tiwari, and K. H. Al-Yahyaee, "An analysis of the weak form efficiency, multifractality and long memory of global, regional and European stock markets," *The Quarterly Review of Economics and Finance*, vol. 72, pp. 168–177, 2019.
- [17] D. Kahneman and A. Tversky, "Prospect theory: an analysis of decision under risk," in *Handbook of the Fundamentals of Financial Decision Making: Part I* World Scientific, Singapore, 2013.
- [18] B. G. Malkiel, "The efficient market hypothesis and its critics," *Journal of Economic Perspectives*, vol. 17, no. 1, pp. 59–82, 2003.
- [19] N. E. Al-Loughani, K. M. Al-Saad, and M. M. Ali, "The holiday effect and stock return in the Kuwait stock exchange," *Global Competitiveness*, vol. 13, no. 1, pp. 81–91, 2005.
- [20] C. Hommes, "Booms, busts and behavioural heterogeneity in stock prices," *Journal of Economic Dynamics and Control*, vol. 80, pp. 101–124, 2015.
- [21] L. J. Cao and F. E. H. Tay, "Support vector machine with adaptive parameters in financial time series forecasting," *IEEE Transactions on Neural Networks*, vol. 14, no. 6, pp. 1506–1518, 2003.
- [22] W. Huang, Y. Nakamori, and S.-Y. Wang, "Forecasting stock market movement direction with support vector machine," *Computers & Operations Research*, vol. 32, no. 10, pp. 2513–2522, 2005.
- [23] C.-J. Lu, T.-S. Lee, and C.-C. Chiu, "Financial time series forecasting using independent component analysis and support vector regression," *Decision Support Systems*, vol. 47, no. 2, pp. 115–125, 2009.
- [24] A. E. Mohamed, "Comparative study of four supervised machine learning techniques for classification," *International Journal of Applied*, vol. 7, no. 2, 2017.
- [25] M. Azimi-Pour, H. Eskandari-Naddaf, and A. Pakzad, "Linear and non-linear SVM prediction for fresh properties and compressive strength of high volume fly ash self-compacting concrete," *Construction and Building Materials*, vol. 230, p. 117021, 2020.
- [26] T. Syriopoulos, M. Tsatsaronis, and I. Karamanos, "Support vector machine algorithms: an application to ship price forecasting," *Computational Economics*, vol. 57, no. 1, pp. 55–87, 2020.
- [27] B. G. Malkiel, "Efficient market hypothesis," in *Finance*-Springer, Berlin, Germany, 1989.
- [28] A. T. M. Shaker, "Testing the weak-form efficiency of the Finnish and Swedish stock markets," *European Journal of Business and Social Sciences*, vol. 2, no. 9, pp. 176–185, 2014.
- [29] A. W. Lo and A. C. MacKinlay, "Stock market prices do not follow random walks: evidence from a simple specification test," *Review of Financial Studies*, vol. 1, no. 1, pp. 41–66, 1988.
- [30] K. V. Chow and K. C. Denning, "A simple multiple variance ratio test," *Journal of Econometrics*, vol. 58, no. 3, pp. 385–401, 1993.
- [31] J. H. Wright, "Alternative variance-ratio tests using ranks and signs," *Journal of Business & Economic Statistics*, vol. 18, no. 1, pp. 1–9, 2000.
- [32] S. Tokić, B. Bolfek, and A. R. Peša, "Testing efficient market hypothesis in developing Eastern European countries," *Investment Management and Financial Innovations*, vol. 15, no. 2, p. 281, 2018.
- [33] J. Franks, R. Harris, and S. Titman, "The postmerger share-price performance of acquiring firms," *Journal of Financial Economics*, vol. 29, no. 1, pp. 81–96, 1991.
- [34] M. M. Cornett and H. Tehranian, "Changes in corporate performance associated with bank acquisitions," *Journal of Financial Economics*, vol. 31, no. 2, pp. 211–234, 1992.
- [35] A. Yalama and S. Çelik, "Financial market efficiency in Turkey: empirical evidence from Toda Yamamoto causality test," *European Journal of Economics, Finance and Administrative Sciences*, vol. 13, pp. 88–93, 2021.
- [36] S. Dharmasena and D. Bessler, "Weak-form efficiency vs semi-strong form efficiency in price discovery: an application to international black tea markets," *Sri Lankan Journal of Agricultural Economics*, vol. 6, no. 1, pp. 1–24, 2015.
- [37] S. S. Ali, K. Mustafa, and A. Zaman, "Testing semi-strong form efficiency of stock market," *The Pakistan Development Review*, pp. 651–674, 2021.
- [38] B. M. Hussin, A. D. Ahmed, and T. C. Ying, "Semi-strong form efficiency: market reaction to dividend and earnings announcements in Malaysian stock exchange," *IUP Journal of Applied Finance*, vol. 16, no. 5, 2010.
- [39] T. Mallikarjunappa and J. J. Dsouza, "A study of semi-strong form of market efficiency of Indian stock market," *Amity Global Business Review*, vol. 8, pp. 60–68, 2017.
- [40] C. O. Manasseh, C. K. Ozuzu, and J. E. Ogbuabor, "Semi strong form efficiency test of the Nigerian stock market: evidence from event study analysis of bonus issues,"



- International Journal of Economics and Financial Issues*, vol. 6, no. 4, 2016.
- [41] A. A. Syed, P. Liu, and S. D. Smith, "The exploitation of inside information at the Wall Street Journal: a test of strong form efficiency," *The Financial Review*, vol. 24, no. 4, pp. 567–579, 1989.
- [42] A. Kara and K. C. Denning, "A model and empirical test of the strong form efficiency of US capital markets: more evidence of insider trading profitability," *Applied Financial Economics*, vol. 8, no. 3, pp. 211–220, 1998.
- [43] T. Potocki and T. Swist, "Empirical test of the strong form efficiency of the Warsaw stock exchange: the analysis of WIG 20 index shares," *South-Eastern Europe Journal of Economics*, vol. 102 pages, 2012.
- [44] D. Han, L. Ma, and C. Yu, "Financial prediction: application of logistic regression with factor analysis," in *Proceedings of the 2008 4th International Conference on Wireless Communications, Networking and Mobile Computing*, pp. 1–4, Dalian, China, October 2008.
- [45] Z. Konglai and L. Jingjing, "Studies of discriminant analysis and logistic regression model application in credit risk for China's listed companies," *Management Science and Engineering*, vol. 4, no. 4, pp. 24–32, 2021.
- [46] S. B. Jabeur, "Bankruptcy prediction using partial least squares logistic regression," *Journal of Retailing and Consumer Services*, vol. 36, pp. 197–202, 2017.
- [47] A. A. Rafatnia, A. Suresh, L. Ramakrishnan, D. F. B. Abdullah, F. M. Nodeh, and M. Farajnezhad, "Financial distress prediction across firms," *Journal of Environmental Treatment Techniques*, vol. 8, no. 2, pp. 646–651, 2020.
- [48] D. Jovanović, M. Todorović, and M. Grbić, "Financial indicators as predictors of illiquidity," *Romanian Journal of Economic Forecasting*, vol. 20, no. 1, pp. 128–149, 2017.
- [49] A. Strzelecka, A. Kurdyś-Kujawska, and D. Zawadzka, "Application of logistic regression models to assess household financial decisions regarding debt," *Procedia computer science*, vol. 176, pp. 3418–3427, 2020.
- [50] V. Vapnik and A. Y. Lerner, "Recognition of patterns with help of generalized portraits," *Avtomat i Telemekh*, vol. 24, no. 6, pp. 774–780, 1963.
- [51] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [52] S. Y. Kim, "Prediction of hotel bankruptcy using support vector machine, artificial neural network, logistic regression, and multivariate discriminant analysis," *Service Industries Journal*, vol. 31, no. 3, pp. 441–468, 2021.
- [53] E. Schöneburg, "Stock price prediction using neural networks: a project report," *Neurocomputing*, vol. 2, no. 1, pp. 17–27, 2020.
- [54] R. Ren, D. D. Wu, and T. Liu, "Forecasting stock market movement direction using sentiment analysis and support vector machine," *IEEE Systems Journal*, vol. 13, no. 1, pp. 760–770, 2018.
- [55] M. Vijh, D. Chandola, V. A. Tikkiwal, and A. Kumar, "Stock closing price prediction using machine learning techniques," *Procedia computer science*, vol. 167, pp. 599–606, 2020.
- [56] M. Qiu and Y. Song, "Predicting the direction of stock market index movement using an optimized artificial neural network model," *PLoS One*, vol. 11, no. 5, p. e0155133, 2016.
- [57] C. Fawson, T. F. Glover, W. Fang, and T. Chang, "The weak-form efficiency of the Taiwan share market," *Applied Economics Letters*, vol. 3, no. 10, pp. 663–667, 1996.
- [58] M. A. Moustafa: Testing the Weak-Form Efficiency of the United Arab Emirates Stock Market.
- [59] K. M. Ahmad, S. Ashraf, and S. Ahmed, "Testing weak form efficiency for Indian stock markets," *Economic and Political Weekly*, vol. 4, pp. 49–56, 2006.
- [60] S. Nisar and M. Hanif, "Testing weak form of efficient market hypothesis: empirical evidence from South Asia," *World Applied Sciences Journal*, vol. 17, no. 4, pp. 414–427, 2012.
- [61] K. Hamid, M. T. Suleman, S. Z. Ali Shah, and R. S. Imdad Akash, "Testing the weak form of efficient market hypothesis: empirical evidence from Asia-Pacific markets," *International Research Journal of Finance and Economics*, vol. 58, no. 2010, 2017.
- [62] W. C. Wei, "Liquidity and market efficiency in cryptocurrencies," *Economics Letters*, vol. 168, pp. 21–24, 2018.
- [63] J. V. Bradley, *Distribution-free Statistical Tests*, Prentice-Hall, New York, NY, USA, 1968.
- [64] W. F. Sharpe, "The sharpe ratio," *The Journal of Portfolio Management*, vol. 21, no. 1, pp. 49–58, 1994.
- [65] F. Zhou, Q. Zhang, D. Sornette, and L. Jiang, "Cascading logistic regression onto gradient boosted decision trees for forecasting and trading stock indices," *Applied Soft Computing*, vol. 84, p. 105747, 2019.
- [66] T. T. Huynh, *Application of Machine Learning in CAPM*, Faculty of Finance, University of Economics Ho Chi Minh City, Chi Minh City, Vietnam, 2020.
- [67] J. Sharma and R. E. Kennedy, "A comparative analysis of stock price behavior on the Bombay, London, and New York stock exchanges," *Journal of Financial and Quantitative Analysis*, vol. 12, no. 3, pp. 391–413, 1997.
- [68] M. A. El-Erian and M. S. Kumar, "Emerging equity markets in Middle Eastern countries," *Staff Papers - International Monetary Fund*, vol. 42, no. 2, pp. 313–343, 1995.
- [69] R. Mookerjee and Q. Yu, "An empirical analysis of the equity markets in China," *Review of Financial Economics*, vol. 8, no. 1, pp. 41–60, 1999.
- [70] N. Groenewold, S. H. K. Tang, and Y. Wu, "The efficiency of the Chinese stock market and the role of the banks," *Journal of Asian Economics*, vol. 14, no. 4, pp. 593–609, 2003.
- [71] D. A. Salazar, J. I. Vélez, and J. C. Salazar, "Comparison between SVM and logistic regression: which one is better to discriminate?" *Revista Colombiana de Estadística*, vol. 35, no. SPE2, pp. 223–237, 2019.
- [72] D. Khanna, R. Sahu, V. Baths, and B. Deshpande, "Comparative study of classification techniques (SVM, logistic regression and neural networks) to predict the prevalence of heart disease," *International Journal of Machine Learning and Computing*, vol. 5, no. 5, pp. 414–419, 2015.
- [73] K. C. Ho, W. Speier, and S. El-Saden, "Predicting discharge mortality after acute ischemic stroke using balanced data," in *Proceedings of the AMIA Annual Symposium Proceedings*, Washington, NJ, USA, November 2014.
- [74] P. Amini, H. Ahmadiania, J. Poorolajal, and M. M. Amiri, "Evaluating the high risk groups for suicide: a comparison of logistic regression, support vector machine, decision tree and artificial neural network," *Iranian Journal of Public Health*, vol. 45, no. 9, p. 1179, 2021.
- [75] K.-J. Kim, "Financial time series forecasting using support vector machines," *Neurocomputing*, vol. 55, no. 1–2, pp. 307–319, 2003.
- [76] Y. Kara, M. Acar Boyacioglu, and Ö. K. Baykan, "Predicting direction of stock price index movement using artificial neural networks and support vector machines: the sample of the Istanbul stock exchange," *Expert Systems with Applications*, vol. 38, no. 5, pp. 5311–5319, 2011.
- [77] J. Patel, S. Shah, P. Thakkar, and K. Kotecha, "Predicting stock and stock price index movement using trend deterministic

- data preparation and machine learning techniques,” *Expert Systems with Applications*, vol. 42, no. 1, pp. 259–268, 2015.
- [78] D. Duong, T. Nguyen, and M. Dang, “Stock market prediction using financial news articles on ho chi minh stock exchange,” in *Proceedings of the 10th International Conference on Ubiquitous Information Management and Communication*, pp. 1–6, Danang, Vietnam, January 2016.
- [79] X. Ji, J. Wang, and Z. Yan, “A stock price prediction method based on deep learning technology,” *International Journal of Crowd Science*, vol. 5, no. 1, pp. 55–72, 2021.



## Retraction

# Retracted: A Game Model for Incentive Mechanism of Distributed Nodes in Supply Chains

### Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] J. Jiang and A. Lyu, "A Game Model for Incentive Mechanism of Distributed Nodes in Supply Chains," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 7502860, 10 pages, 2021.

## Research Article

# A Game Model for Incentive Mechanism of Distributed Nodes in Supply Chains

Jingjing Jiang <sup>1</sup> and Aobo Lyu <sup>2</sup>

<sup>1</sup>School of Business, Beijing Union University, Beijing, China

<sup>2</sup>School of Public Administration, Sichuan University, Chengdu, China

Correspondence should be addressed to Jingjing Jiang; [jiangjingruc@163.com](mailto:jiangjingruc@163.com)

Received 12 October 2021; Accepted 3 November 2021; Published 7 December 2021

Academic Editor: Daqing Gong

Copyright © 2021 Jingjing Jiang and Aobo Lyu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This study aims to solve the credit problems in the supply chain commodity and currency circulation links from the perspective of the ledger, while the game model method has been adopted. The research firstly reviews the relationship between distributed ledger technology and the essential functions of currency. Then, by constructing two-agent single-period and multi-period game models in the entire supply chain, the researchers analysed the incentive mechanism and equilibrium solution of distributed nodes of Central Bank Digital Currency (CBDC). The results of this study include the incentive mechanism and optimization of distributed nodes based on licensed distributed ledger technology, which is an important issue that CBDC faces when performing currency functions. The implications of this study mainly cover the limitations of the underlying technology of the public chain and its reward mechanism in the supply chain management and provide support for the rationality of the CBDC issuance mechanism based on state-owned commercial banks, which provides a reference for the CBDC practice. The main value of the research not only serves the decision-making department of the CBDC issuance but also provides ideas on the operation mode of digital currency for the field of digital currency research.

## 1. Introduction

At present, the global macroenvironment is complex and severe, economic growth has fallen, commodity prices have risen, epidemic trends and economic trends have become complex, and risks and challenges continue. In this context, complying with the digital development trend and doing an excellent job in commodity quality management in a long-term and reliable way has become an essential topic in supply chain management.

Commodity quality management and credit issues have long been concerns in the research field of supply chain management. Although studies have shown that information sharing between supply chains can significantly improve product quality and business performance [1], and producers can contract design and revenue distribution mechanisms to avoid immoral behavior of suppliers [2]; there is still a gap in the research on the authenticity guarantee of credit and commodity quality at the technical

level. Some studies have pointed out that blockchain is a new and revolutionary technology that significantly affects the supply chain network and discussed the significance of blockchain technology in supply chain management [3, 4]. Therefore, nowadays, with the innovation and development of distributed ledger technology and its application in Central Bank Digital Currency (CBDC) research and development, using CBDC to support information sharing activity and solving credit and commodity quality management problems in the supply chain scenario has gradually become possible.

Distributed ledger technology can express rules through algorithm programs to enable participants to trust standard algorithm programs and establish mutual trust in the R & D and application of CBDC to solve the trust problem of commodity quality management in the supply chain scenario with higher efficiency and lower cost. This study intends to take this as the research goal, build a game model to simulate the incentive selection of distributed nodes in

distributed ledger technology, and obtain the optimal trusted manager of commodity quality by analyzing the equilibrium solution, to provide a reference for the application of distributed ledger technology in supply chain management and CBDC practice.

The follow-up structure of this paper is organized as follows. Section 1 discusses the theoretical basis for the possible realization of the function of digital currency by distributed ledger technology and briefly describes its optimization mechanism for the existing currency and how to use distributed ledger technology to realize the function of currency and what problems to be solved urgently. Section 2 analyses digital currency's incentive and restraint mechanism based on distributed ledger technology by constructing a game model. Section 3 attempts to solve the equilibrium conditions and puts forward the optimal digital currency distributed nodes selection. Section 4 is the conclusion of this paper.

## 2. Related Works

*2.1. Overview of Distributed Ledger Technology.* Distributed ledger technology is based on cryptography. Designing and implementing the "consensus" mechanism in multiple distributed nodes can completely and accurately record the complete historical process of transactions and payments. The ledger based on the distributed ledger technology is a publicly visible decentralized shared ledger. "Consensus" is a necessary condition for the distributed ledger to achieve the consistency of multinode records, the robustness of consensus protocols, the efficiency of reaching a steady state through high-speed convergence, and the security of the system [5]. From the technical perspective, the currently recognized consensus mechanisms include proof of work (POW), proof of stake (POS), PBFT, RSCoin, hybrid consensus, and consensus mechanism based on DAG technology.

*2.2. Currency Function under Licensed Distributed Ledger Technology.* Under the licensed distributed ledger technology, each transaction and payment is updated to the ledger by the licensed distributed node. A consensus is reached through confirmation if it is confirmed as a real transaction or payment by an absolute majority. Among them, how to design the absolute majority protocol to ensure that the distributed node data changing with the transaction and payment is confirmed by the absolute majority protocol and finally reach an agreement, that is, the "consensus" mechanism. At present, it has been used in security settlement systems, trade settlement systems, central bank digital currency design, and "stable currency."

In a single transaction and payment, the specific process is as follows: different licensed distributed nodes update the ledger according to their respective transactions and payments. Each update needs to be confirmed by an absolute majority of agreements. If it is confirmed as a real transaction and payment, it will be updated in the decentralized shared ledger (DSL). The working principle of the licensed distributed ledger technology is shown in Figure 1.

In society's transactions and payments, smart contracts will record the transactions of goods and services and the receipt and payment process of funds. The use of smart contracts facilitates the recording and verification of distributed ledgers. For manufacturers that generate inferior products and provide inferior services mixed in the transaction and payment process, most nodes of the licensed distributed ledger will verify and update the information and provide the information to the fund payer for decision-making. The working principle of licensed distributed ledger technology in social transactions and payments is shown in Figure 2.

The main role of distributed ledger technology in performing monetary functions is to prevent "double flowers." All transactions and payments, either using traditional online or offline modes or equipped with smart contracts, can be verified by licensed nodes using distributed ledger technology to avoid the "double flower" problem. The transactions and payments under the distributed ledger technology are recorded through cryptography technology. We can know whether the historical records have been tampered with by encryption and decryption. If the absolute majority is verified to be "true" by the licensed distributed nodes, it proves that the digital currency has not been a "double flower," and the transactions and payments are effective.

### 2.3. Problems Faced by Digital Currency Performing Monetary Function Based on Distributed Ledger Technology

*2.3.1. How to Select Distributed Nodes?* Under the unlicensed distributed ledger technology, any willing and capable economic individual can act as a distributed node to update and verify the ledger. The current public chain chooses to use the unlicensed distributed ledger technology. However, the demanders for digital currency involve almost all economic individuals, and the number of transactions and payments is vast. Each transaction and payment need to be confirmed based on the absolute majority of the consensus mechanism. It takes a lot of computing power under the distributed ledger technology without a license. At the same time, considering the rapid layout and development of quantum technology, the tamper-proof function based on a large amount of computing power investment is at risk of being cracked. Therefore, digital currencies with high-security requirements should adopt licensed distributed ledger technology.

However, even with licensed distributed ledger technology, digital currency still faces the problem of distributed node effectiveness in performing its currency function. Effective distributed nodes need incentives to provide real authentication information to achieve antitampering and avoid "double flowers." For licensed distributed nodes, it is difficult to technically ensure that they fully participate in the verification and signature of each transaction, and it is also difficult to technically prohibit them from verifying multiple account books with conflicting accounting items. Because of this, how to select the desired distributed nodes to correctly select and verify the "right" ledger and the "wrong" ledger?

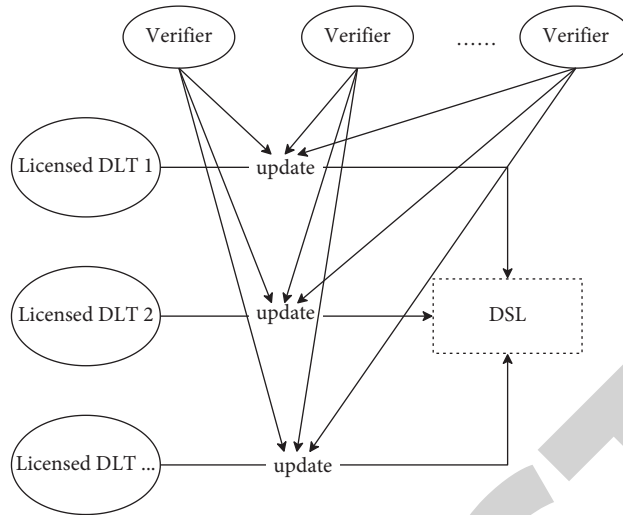


FIGURE 1: How the licensed distributed ledger technology works?

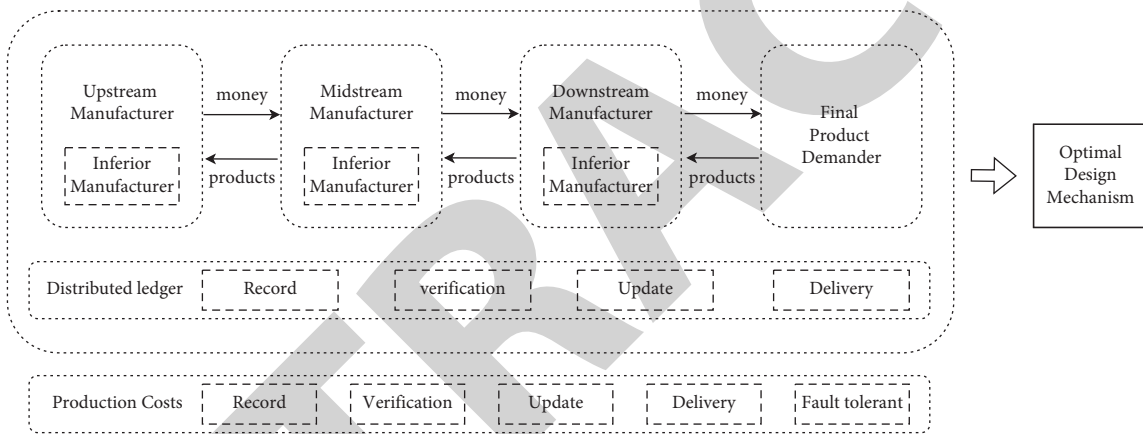


FIGURE 2: How the licensed distributed ledger technology works in social transactions and payments?

2.3.2. *How to Design Incentive Mechanism?* How to design an incentive mechanism to ensure the integrity and authenticity of verification for acceptable licensed distributed nodes? The selected distributed nodes can provide real authentication information and ensure that each transaction can be verified. In addition, how to design a better incentive mechanism so that the verification process of distributed nodes can promote the improvement of transaction scale and quality in the economy?

In this regard, Nosal and Rocheteau [6] proposed that if distributed nodes can publicly obtain historical transaction data, they will have less motivation to make false verification. At the same time, due to reducing the verification cost, it may improve the integrity of verification. Furthermore, the disclosure of historical transactions can help promote the completion of transactions to promote economic growth and high-quality development. However, historical transaction data cannot be obtained completely and in real time [7]. How to design an effective incentive mechanism under the condition of incomplete information so that the income obtained by the information verifier is greater than the cost so that the distributed nodes can provide “pair” verification

and verify each transaction and improve the transaction scale and quality at the same time? What is included in the incentive mechanism?

### 3. Methods

3.1. *Model Overview.* The unlicensed distributed ledger technology may be more vulnerable to historical tampering attacks based on the above analysis. Therefore, the model constructed in this paper should be based on licensed distributed ledger technology. At the same time, under the assumption of incomplete information, an incentive mechanism should be constructed to enable the desired distributed nodes to provide “right” verification as the verifier of transaction information to verify each transaction and improve the decision-making of transaction scale and quality. Therefore, the incentive mechanism should enable the verifier’s “verification of” to be partially observed and rewarded, and vice versa. In this regard, Carlsson and Damme [8] proposed the research methods of multiagent information incomplete multistage game and global game, which is the modelling method of this research. It is assumed

that decision makers can only observe the profit and loss with continuous noise in the game in the economy with incomplete information, and the equilibrium result is unique, which is a research method with high consistency with the research purpose of this paper.

Morris and Shin [9–11] have used this method to make a series of studies on the self-realization mechanism of the financial crisis and put forward policy suggestions according to the equilibrium results. Brown et al. [12], Hakenes and Schliephake [13], He and Manela [14], Schilling [15], and Yang et al. [4] all used game analysis methods to study the liquidity and system stability in the financial field. Some studies have also used game theory to study the interaction between suppliers and sellers in the supply chain [16, 17]. Therefore, this paper intends to build a game model on this basis and design the digital currency transaction verification incentive mechanism and consensus mechanism based on the licensed distributed ledger technology according to the equilibrium results of the model.

**3.2. Model Assumptions.** Based on the working principle of licensed distributed ledger technology in social transactions and payments shown in Figure 2, this study establishes a model to analyze the selection of producers and verifiers. The operation of the model follows the following assumptions.

**Assumptions 1.** The whole social production is divided into two stages, and consumers can only obtain utility from the final products or services provided by downstream manufacturers. The producers are divided into stage-I producers and stage-II producers. The former produces raw materials, including upstream and midstream manufacturers in Figure 2, and the stage-II producers produce final products, i.e., downstream manufacturers in Figure 2.

**Assumptions 2.** In the whole society, consumers are producers. Then, consumers include stage-I producers and stage-II producers, that is, the demander of the final product in Figure 2.

**Assumptions 3.** Inferior producers and inferior service providers only exist in stage-II producers. This is because the stage-II producers produce the final products, and the stage-II producers consume the final products. Therefore, for the stage-II producers, this group consumes the goods produced or services they provide. Therefore, it is assumed that there is no possibility of producing inferior products or providing inferior services.

**Assumptions 4.** Based on the licensed distributed ledger technology, since only stage-I producer may choose to produce inferior goods or provide inferior services, only stage-I producer may be involved in whether the confirmed transaction and payment information are “true.” Therefore, it is assumed that the confirmation of whether the transaction and payment information is “true” can only be completed by stage-II producers.

**Assumptions 5.** The transaction and payment system contain several cycles. In the first production cycle, if the stage-II producer provides false information as the information confirmer, it will be punished by being excluded from the transaction and payment system. Therefore, it is assumed that the stage-II producer will not provide false information in the first production cycle.

**Assumptions 6.** Each entity makes the current strategy choice according to the historical decisions of other entities.

### 3.3. Model Development

**3.3.1. Utility Analysis of Various Subjects.** If the proportion of manufacturers’ producing inferior products and providing inferior services in the stage I is  $f$  ( $f$  stands for fault), the total number of manufacturers in the stage II becomes  $1 - f$  (this is because the manufacturers producing inferior products and providing inferior services in the stage I are observed, so they are excluded from the stage-II trading system) of the total number in stage I.

**(1) Stage-I Qualified Producers.** The utility of a qualified producer in stage I comes from the difference between the goods or services enjoyed and the goods or services provided. Therefore, the utility function of a qualified producer in stage I is

$$U_1 = U_1[(1 - f)y_2] - P(1 - f)y_2 - y_1 + Py_1, \quad (1)$$

where  $U_1$  represents the utility function of qualified producers in stage I,  $y_1$  refers to the qualified goods or services produced and sold by producers in stage I, that is, the qualified goods or services consumed by producers in stage II,  $y_2$  means the qualified goods or services produced and sold by the stage-II producer, that is, the qualified goods or services consumed by the stage-I producer, then  $(1 - f)y_2$  refers to the qualified goods or services consumed by qualified producers in stage I,  $U_1[(1 - f)y_2]$  refers to the utility obtained by a qualified producer in stage I from the qualified goods or services consumed by him,  $P$  represents the general price level of goods or services,  $Py_1$  represents the production income of qualified producers in stage I, and  $P(1 - f)y_2$  represents the consumption expenditure of qualified producers in stage I.

**(2) Inferior Producers and Inferior Service Providers in Stage I.** The utility of inferior producers and service providers in stage I comes from the qualified goods or services in stage II. Therefore, the utility function of inferior producers and service providers in stage I is

$$U_{1f} = U_1(fy_2) - Pfy_2, \quad (2)$$

where  $fy_2$  refers to the qualified goods or services consumed by inferior producers in stage I,  $U_1(fy_2)$  refers to the utility obtained by inferior producers in stage I from the qualified goods or services consumed by them, and  $Pfy_2$  represents the consumption expenditure of inferior producers in stage I.

(3) *Stage-II Qualified Producers.* The utility of the stage-II producer comes from the difference between the utility enjoyed by the qualified goods or services purchased from stage I producer as raw materials and the goods or services provided. Therefore, the utility function of the stage-I producer is

$$U_2 = U_2(y_1) - Py_1 - y_2 + Py_2, \quad (3)$$

where  $U_2(y_1)$  refers to the utility enjoyed by the stage-II producer in purchasing qualified goods or services as raw materials from the stage-I producer and  $Py_2$  represents the production income of stage-II producers.

### 3.3.2. Game Strategy Selection of Various Subjects

(1) *Stage-I Producers.* At stage I, producers can choose two strategies: one is to produce qualified products or provide qualified services, and the other is to produce inferior products or provide inferior services.

(2) *Stage-II Producers.* According to the analysis of the hypothesis part of this paper, the stage-II producers can only produce qualified products or provide qualified services, so there is no strategy selection in production. As shown in Figure 2, the strategy selection of stage-II producers may occur in the information confirmation stage based on licensed distributed ledger technology. At the same time, according to the analysis of the hypothesis part of this paper, only stage-I producers may choose to produce inferior products or provide inferior services. Therefore, only stage-I producers may involve the confirmed information, which means the stage-II producers can choose two strategies: one is confirmed as “true,” and the other is not confirmed as “true.”

### 3.3.3. Game Profit and Loss Matrix of Various Subjects

(1) *Profit and Loss Matrix in the First Complete Production Cycle.* In the first production cycle, various entities do not master the historical strategy choices of other entities; especially, the stage-II producers, who undertake the task of information confirmation, are likely to make wrong decisions in the first production cycle. According to the utility analysis of various entities, the following four quadrant profit and loss matrix is obtained, as shown in Table 1.

**Profit and Loss Analysis of Each Entity in the First Quadrant.** If all stage-I producers choose to produce qualified products or provide qualified services, and stage-II producers, as information confirmers, make correct judgment and confirm it as “true,” the profit and loss obtained by stage-II producers is as follows.  $U_2(y_1) - Py_1 - y_2 + Py_2 + R$  refers to the utility enjoyed by qualified goods or services purchased as raw materials from stage-I producers, minus the consumption expenditure paid for this, then excluding the payment for stage-II production, plus the income obtained therefrom and finally plus the remuneration obtained

as an information confirmer. The profit and loss obtained by the producer in the corresponding stage I is  $U_1(y_2) - Py_2 - y_1 + Py_1$ , that is, the utility enjoyed by qualified goods or services purchased from stage-II producers, minus the consumer expenditure paid for this, and then excluding the payment for stage-I production, plus the income obtained therefrom.

**Profit and Loss Analysis of Each Entity in the Second Quadrant.** If stage-I producer chooses to produce inferior products or provide inferior services, and the stage-II producer, as the information confirmer, makes a wrong judgment and confirms it as “true,” the profit and loss obtained by the stage-II producer is  $U_2[(1-f)y_1] - Py_1 - y_2 + Py_2 + R$ , which refers to the utility enjoyed by qualified goods or services purchased as raw materials from stage-I producers minus the consumption expenditure paid for it. Since stage-II producers can only obtain utility from qualified products or services provided by stage-I producers, but do not confirm inferior products, they can only enjoy the utility of  $1-f$  ratio. However, it is necessary to pay for all products or services (including inferior products and services), exclude the payment for the stage-II production, add the income thus obtained, and finally add the remuneration obtained as the information confirmer. The profit and loss obtained by the producer in the corresponding stage I is  $U_1(fy_2) - Pfy_2 + Py_1$ , that is, the utility enjoyed by qualified goods or services purchased from stage-II producers, minus the consumer expenditure paid for this, plus the income from the provision of all products or services (including inferior goods and inferior services).

**Profit and Loss Analysis of Three Quadrant Entities.** If all stage-I producers choose to produce qualified products or provide qualified services, but stage-II producers, as information confirmers, identify the part with proportion  $F$  as inferior product producers or inferior service providers, and therefore do not confirm it as “true,” the profit and loss obtained by stage II producers is  $U_2[(1-f)y_1] - P(1-f)y_1 - y_2 + Py_2 + (1-f)R$ , which refers to the utility enjoyed by qualified goods or services purchased from stage-I producers as raw materials, minus the consumption expenditure paid for this, and then excluding the payment for stage-II production, plus the income obtained therefrom. It should be explained that based on the licensed distributed ledger technology if the information is determined not to be “true,” it does not need to be confirmed, so it is impossible to obtain the remuneration of the information confirmer. Therefore, in this case, the stage-II producer only obtains the corresponding confirmation remuneration for the part confirmed as “true.” The profit and loss obtained by the producer in the corresponding stage I is  $U_1(y_2) - Py_2 - y_1 + P(1-f)y_1$ , that is, the utility enjoyed by qualified goods or services purchased from stage-II producers, minus the consumer expenditure paid for this, excluding the payment for stage-I production, plus the income from products or services confirmed as “real.”

**Profit and Loss Analysis of Four Quadrant Entities.** If stage-I producer chooses to produce inferior products or



TABLE 1: Profit and loss matrix of each entity in the first complete production cycle.

		Stage-I producers	
		Produce qualified products	Produce inferior products
Stage-II producers	Confirm it as "true"	$U_2(y_1) - Py_1 - y_2 + Py_2 + R$ and $U_1(y_2) - Py_2 - y_1 + Py_1$	$U_2[(1-f)y_1] - Py_1 - y_2 + Py_2 + R$ and $U_1(fy_2) - Pfy_2 + Py_1$
	Not confirm it as "true"	$U_2[(1-f)y_1] - P(1-f)y_1 - y_2 + Py_2 + (1-f)R$ and $U_1(y_2) - Py_2 - y_1 + P(1-f)y_1$	$U_2[(1-f)y_1] - P(1-f)y_1 - y_2 + Py_2 + (1-f)R$ and $U_1(fy_2) - Pfy_2$

provide inferior services and the stage-II producer, as the information confirmer, makes a correct judgment and does not confirm it as "true," the profit and loss obtained by the stage II producer is  $U_2[(1-f)y_1] - P(1-f)y_1 - y_2 + Py_2 + (1-f)R$ , i.e., the utility enjoyed by the qualified goods or services purchased as raw materials from the producers in stage I, minus the consumption expenditure paid for this purpose and plus the recognition remuneration obtained for the part confirmed as "true." The profit and loss obtained by the producer in the corresponding stage I is  $U_1(fy_2) - Pfy_2$ , that is, the utility enjoyed by qualified goods or services purchased from stage-II producers minus the consumer expenditure paid for this.

(2) *Profit and Loss Matrix in Subsequent Production Cycles.* Under the assumption of adaptive expectation, the stage-II producer as the information confirmer can obtain feedback according to the consumption feeling after the first production cycle and identify and eliminate the inferior producers and inferior service providers from the trading and payment system, so as to adjust the information confirmation in the subsequent production cycle. Based on this, the stage-II producer as the information confirmer may make a wrong judgment because stage-I producer no longer provides qualified products or services and chooses to produce inferior products or provide inferior services. Based on this, the following four-image limit profit and loss matrix for the subsequent production cycle is obtained, as shown in Table 2.

**Profit and Loss Analysis of Each Entity in the First Quadrant.** If all the stage-I producers choose to produce qualified products or provide qualified services and the stage-II producers, as the information confirmer, make a correct judgment and confirm it as "true," the profits and losses obtained by the stage-II producers are the same as those in the first complete production cycle and will not change.

**Profit and Loss Analysis of Each Entity in the Second Quadrant.** In the first production cycle, the stage-II producers did not exclude the stage-I producers who chose to provide inferior products or services from the trading and payment system, but they were identified after consumption experience. The "missed fish" can only be lucky once. Therefore, in the subsequent production cycle, it is only possible for the stage-I producers to choose to provide inferior products or services, At the same time, the stage-II producer, as the information confirmer, makes a wrong judgment and confirms it as "true." Therefore, the profits and losses of producers in stage I and stage II of each production cycle remain unchanged, but the individual producers in stage I who choose to provide inferior products or services have changed.

**Profit and Loss Analysis of Three Quadrant Entities.** The profit and loss of the three quadrants in the subsequent production cycle are similar to that of the two quadrants. The misjudgment made by the stage-II producer in the first production cycle will be identified afterwards through the consumption experience, and the misjudgment will only occur once. Therefore, in the subsequent production cycle, it can only be a new misjudgment, and the stage-I producer providing qualified products or services is not confirmed as "true." Therefore, the profits and losses obtained by producers in stage I and stage II of each production cycle remain unchanged, but the misjudged individuals have changed.

**Profit and Loss Analysis of Four Quadrant Entities.** As the stage-II producers in each production cycle correctly identify the stage-I producers providing inferior products or services and remove them from the trading and payment system, there should be fewer and fewer stage-I producers actively choosing to provide inferior products or services; then, the profit and loss obtained by the stage-II producers is  $U_2[(1-f)^n y_1] - P(1-f)^n y_1 - y_2 + Py_2 + (1-f)R$ ; the profit and loss obtained by the producer in the corresponding stage I is  $U_1(f^n y_2) - P f^n y_2$ .

## 4. Results and Discussion

### 4.1. Equilibrium Solution in the First Complete Production Cycle

**4.1.1. When Stage-I Producer Provides Qualified Products or Services.** For the stage-II producer as the information confirmer, when the stage-I producer chooses to provide qualified products or services, the stage-II producer will choose to make a correct judgment and confirm the production, payment, and transaction information of the stage-I producer as "true." This is because, from the profit and loss matrix, when  $U_2(fy_1) - Pfy_1 + fR > 0$ , the benefit obtained by the stage-II producer when it is confirmed as "true" is greater, and the utility obtained by the stage-II producer from the stage-I producer must be greater than the expenditure, so  $U_2(fy_1) > Pfy_1$ . At the same time, the stage-II producer misjudgment rate  $f > 0$ , and the information confirmation reward  $R > 0$ .

Therefore, when the stage-I producer chooses to provide qualified products or services, the stage-II producer will choose to make a correct judgment and confirm the production, payment, and transaction information of the stage-I producer as "true."

TABLE 2: Profit and loss matrix of each entity in subsequent production cycles.

		Stage-I producers	
		Produce qualified products	Produce inferior products
Stage-II producers	Confirm it as "true"	$U_2(y_1) - Py_1 - y_2 + Py_2 + R$ and $U_1(y_2) - Py_2 - y_1 + Py_1$	$U_2[(1-f)y_1] - Py_1 - y_2 + Py_2 + R$ and $U_1(fy_2) - Pfy_2 + Py_1$
	Not confirm it as "true"	$U_2[(1-f)y_1] - P(1-f)y_1 - y_2 + Py_2 + (1-f)R$ and $U_1(y_2) - Py_2 - y_1 + P(1-f)y_1$	$U_2[(1-f)^n y_1] - P(1-f)^n y_1 - y_2 + Py_2 + (1-f)R$ and $U_1(f^n y_2) - Pf^n y_2$

*4.1.2. When Some Stage-I Producers Provide Inferior Products or Services.* When some stage-I producers provide inferior products or services, whether the stage-II producers can choose to make a correct judgment and not confirm the production, payment, and transaction information of the stage-I producers as "true" depends on the information available to confirm the reward  $R$ . This is because it can be obtained from the profit and loss matrix when  $Pfy_1 - fR > 0$ ; the income obtained by the stage-II producer when it is not confirmed as "true" is greater, that is, when  $R$  is smaller, the stage-II producer is more likely to make a correct judgment. However,  $R$  is the incentive means for stage-II producers as information confirmers. If  $R$  is too small, stage-II producers lack the incentive to confirm information.

Therefore, when some stage-I producers provide inferior products or services, stage-II producers may not be willing to become information confirmers, or it is difficult to make a correct judgment, and the production, payment, and transaction information of stage-I producers providing inferior products or services are not confirmed as "true."

*4.1.3. When the Stage-II Producer Confirms the Information as "True".* When the stage-II producer confirms the information as "true," the stage-I producer will choose to provide inferior products or services. This is because, from the profit and loss matrix, when  $U_1(y_2) - Py_2 - [U_1(fy_2) - Pfy_2] - y_1 < 0$ , the income of producers providing inferior products or services in stage I is greater. According to the principle of diminishing marginal utility, the difference between the utility and expenditure of producers in stage I decrease with the increase of consumption, and the above inequality is true.

*4.1.4. When the Stage-II Producer Does Not Confirm the Information as "True".* When the stage-II producer does not confirm the information as "true," whether the stage I producer chooses to provide inferior products or services depends on the rate  $f$  that the stage-II producer thinks it provides inferior products. This is because, from the profit and loss matrix, when  $U_1(y_2) - Py_2 - [U_1(fy_2) - Pfy_2] - y_1 + P(1-f)y_1 < 0$ , the income of producers providing inferior products or services in stage I is greater. According to the principle of diminishing marginal utility, the difference between the utility and expenditure of producers in stage I decreases with the increase of consumption. At the same time, the greater  $F$ , the greater the possibility of the above inequality.

That is, failure to confirm the information as "true" will bring incentives for producers in stage I to provide inferior products or services.

*4.1.5. Analysis of Equilibrium Solution in the First Complete Production Cycle.* According to the above analysis, based on Hypothesis 4, the stage-II producers should be given a certain information confirmation reward as an incentive. At this time, confirmed as "true," providing inferior products or services is the Nash equilibrium of the game matrix, and the profit and loss combination at this time is  $U_2[(1-f)y_1] - Py_1 - y_2 + Py_2 + R$  and  $U_1(fy_2) - Pfy_2 + Py_1$ .

Under the Nash equilibrium, the stage-II producers who only consider the current income have more incentive to confirm the information as "true" and obtain corresponding remuneration. The stage-I producers who only consider the current income will partially tend to provide inferior products or services, so as to enjoy the income from selling products or providing services, but pay less costs. It can be seen that the incentive mechanism should be optimized for the short-sighted behavior of various subjects, and the long-term equilibrium should be solved by constructing a multiproduction cycle game model.

## 4.2. Equilibrium Solution in Subsequent Production Cycles

*4.2.1. When Stage-I Producer Provides Qualified Products or Services.* When the producers in stage I choose to provide qualified products or services, the profit and loss matrix is in the same form as the first production cycle, so the equilibrium point is also the same. The producers in stage II will choose to make a correct judgment and confirm the production, payment, and transaction information of the producers in stage I as "true."

That is, when the stage-I producer chooses to provide qualified products or services, the stage-II producer will choose to make a correct judgment and confirm the production, payment, and transaction information of the stage-I producer as "true."

*4.2.2. When Some Stage-I Producers Provide Inferior Products or Services.* When some stage-I producers provide inferior products or services, whether the stage-II producers can choose to make a correct judgment and not confirm the production, payment, and transaction information of the stage-I producers as "true" depends on the ratio  $F$  of the stage-I producers providing inferior products or services and the

information confirmation reward  $r$  available. This is because, from the profit and loss matrix, when  $U_2[(1-f)^n y_1] - U_2[(1-f)y_1] - P(1-f)^n y_1 + P y_1 - fR > 0$ , the income obtained by the stage-II producer is greater when it is not confirmed as “true,” that is, when  $f$  and  $R$  are small, the stage-II producer is more likely to make a correct judgment. However, similarly, the size of  $f$  is difficult to control, and too small  $R$  will lead to the lack of incentive of confirmation information for stage-II producers.

Therefore, when some stage-I producers provide inferior products or services, stage-II producers may not be willing to become information confirmers, or it is difficult to make a correct judgment, and the production, payment, and transaction information of stage-I producers providing inferior products or services are not confirmed as “true.”

*4.2.3. When the Stage-II Producer Confirms the Information as “True”.* When the stage-II producer confirms the information as “true,” the profit and loss matrix is the same in form as the first production cycle, so the equilibrium point is also the same. The stage-I producer must choose to provide inferior products or services.

*4.2.4. When the Stage-II Producer Does Not Confirm the Information as “True”.* When the stage-II producer does not confirm the information as “true,” whether the stage-I producer chooses to provide inferior products or services depends on how many production cycles in the stage-II producer acts as the information confirmer and the ratio  $f$  that the stage-II producer believes it provides inferior products. This is because, as can be seen from the profit and loss matrix, if the stage-II producer acts as the information confirmer for a long time, there may be insufficient consumption in the economy due to the elimination of too many stage-I producers in the transaction and payment system;  $U_1(y_2) - P y_2 - [U_1(f^n y_2) - P f^n y_2]$  may be negative first and then positive, but with the increase of the rate  $f$  that stage-II producers think stage-I producers provide inferior products, the value of  $-y_1 + P(1-f)y_1$  will also decrease; then,  $U_1(y_2) - P y_2 - [U_1(f^n y_2) - P f^n y_2] - y_1 + P(1-f)y_1 < 0$  is more likely to be established.

That is, failure to confirm the information as “true” and acting as the information confirmer for a long time will bring incentives for producers in stage I to provide inferior products or services.

*4.2.5. Analysis of Equilibrium Solution in Subsequent Production Cycle.* According to the above analysis, confirmed as “true,” providing inferior products or services is the Nash equilibrium of the game matrix, and the profit and loss

combination at this time is  $U_2[(1-f)y_1] - P y_1 - y_2 + P y_2 + R$  and  $U_1(f y_2) - P f y_2 + P y_1$ .

Under the Nash equilibrium, the stage-II producers have more incentive to confirm the information as “true.” As they act as information confirmers in multiple production cycles, the stage-II producers have more incentive to confirm the transactions and payments of “right.” Of course, at the same time, the producers of false transactions and payments are excluded, and they will be rewarded if they are confirmed as “true.” In stage I, producers will partially tend to provide inferior products or services, so as to enjoy the income from selling products or providing services, but pay less costs.

*4.3. Better Selection of Distributed Nodes.* From the above multiperiod game analysis and equilibrium results of incomplete information, it can be seen that, from the perspective of individual participants, if the stage-II producer acts as the information confirmer, it will face the contradiction of whether it acts as the information confirmer for multiple consecutive periods: if it acts as the information confirmer only in a few production cycles, it is possible to confirm too much as “true” to obtain remuneration. If you act as an information confirmer in consecutive production cycles, there is a risk of insufficient consumption in the economy. At the same time, producers at stage I are also vulnerable to incentives to provide inferior products or services.

Therefore, a better choice is to try to use trusted financial intermediaries as licensed distributed nodes. Trusted financial institutions represented by systemically important banks have a clear motivation to identify producers who provide inferior products or services and can also avoid the problem of false confirmation as “true” in order to obtain information confirmation remuneration. Under this mechanism, the stage-II producer will no longer act as the information confirmer, but trusted financial institutions such as commercial banks that have more information about the manufacturer will assume the function of confirming the information. The transaction information can be confirmed in the form of associated smart contracts, and the transaction information can be confirmed as “true” if it is consistent with the bank account. The optimization scheme of digital currency distributed nodes based on permitted DLT is shown in Figure 3.

## 5. Conclusion

The relationship between currency and credit is inseparable. Currency must have a good credit foundation. At the same time, currency performs its essential functions relying on credit. With the innovative development of distributed ledger technology, the requirements of currency on credit

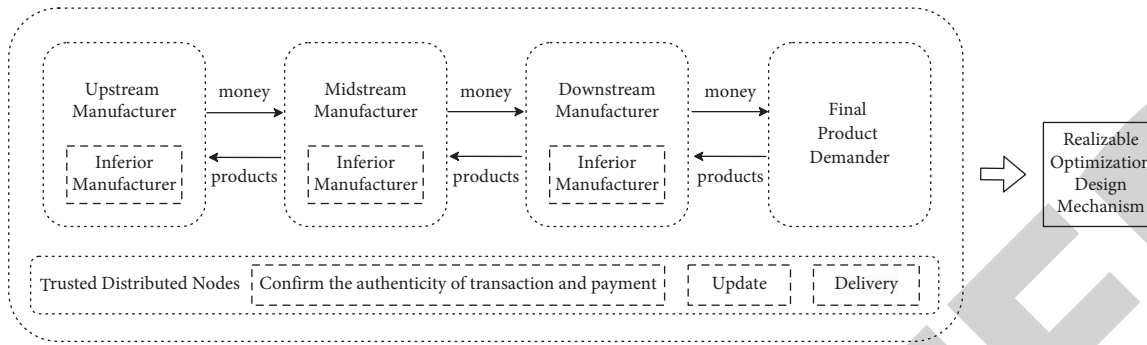


FIGURE 3: Digital currency distributed node optimization scheme based on permitted DLT.

may be further met. This paper explores the close relationship between currency, credit, and account book and provides a theoretical explanation for solving the credit problem in currency from the perspective of account book. Based on the highly consistent relationship between distributed ledger technology and the essential functions of currency, this paper analyses the incentive mechanism of distributed nodes of digital currency. It obtains the equilibrium solution by constructing two main body single-stage and multistage game models. This study found that if the producer of the final product or the provider of the final service is the most licensed distributed node, it may lead to the risk of over recognition and insufficient consumption in the economy. On this basis, aiming at the above possible problems and combined with the reality of China's digital currency issuance, this paper puts forward the optimal selection of digital currency distributed nodes. It uses trusted financial intermediaries as licensed distributed nodes, providing a reference basis for the research and development of China's digital currency and its practice in economy and finance.

Therefore, this paper has some contributions to the research in the field of CBDC. Bis [11] and Xiao (2020) have proposed that since CBDC involves huge transaction and payment data, the distributed ledger technology without a license is not applicable when selecting distributed nodes. For this, Nosal and Rocheteau [6] have demonstrated that licensed distributed ledger technology can improve verification integrity. However, for licensed distributed nodes, it is difficult to technically prohibit them from verifying multiple ledgers with conflicting accounting items. This paper agrees with this view and, on this basis, further discusses how to select the distributed node of CBDC to ensure that it can perform the monetary function well and take into account the cost problem, that is, how to select the desired distributed node so that it can correctly select the account book that verifies "right" and the account book that does not verify "wrong."

In addition, there are still some limitations of this study. This paper puts forward the viewpoint of taking the systemically important commercial bank as the optimally distributed node. This viewpoint is based on the conclusion that the final product producer or the final service provider cannot be the optimally distributed node. Therefore, it still lacks rigorous mathematical inference, which will also be the

research direction and focus in the next step. Moreover, because digital currency has not been widely used in any country or region, the conclusion of this paper also lacks the support of empirical data, which will be supplemented in the subsequent stage of research with the continuous application of digital currency.

### Data Availability

The data used to support the findings of the study are available from the corresponding author upon request.

### Conflicts of Interest

The authors declare that they have no conflicts of interest to declare and state that (i) no support, financial or otherwise, has been received from any organization that may have an interest in the submitted work and (ii) there are no other relationships or activities that could appear to have influenced the submitted work.

### Acknowledgments

This work was supported by Fundamental Research Funds of ICBR (research on collaborative innovation for blockchain-based supply chain in bamboo and rattan industry, Grant no. 1632021028).

### References

- [1] T. H. Nazifa and K. K. Ramachandran, "Information sharing in supply chain management: a case study between the cooperative partners in manufacturing industry," *Journal of System and Management Sciences*, vol. 9, no. 1, pp. 19–47, 2019.
- [2] S. B. Wu, X. Gu, G. D. Wu, and Q. Zhou, "Cooperative R&D contract of supply chain considering the quality of product innovation," *International Journal of Simulation Modelling*, vol. 15, no. 2, pp. 341–351, 2016.
- [3] R. Goyat, G. Kumar, M. K. Rai, and R. Saha, "Implications of blockchain technology in supply chain management," *Journal of System and Management Sciences*, vol. 9, no. 3, pp. 92–103, 2019.
- [4] J. Q. Yang, X. M. Zhang, H. Y. Zhang, and C. Liu, "Cooperative inventory strategy in a dual-channel supply chain with transshipment consideration," *International Journal of Simulation Modelling*, vol. 15, no. 2, pp. 365–376, 2016.

## *Retraction*

# **Retracted: Modeling on Social Health Literacy Level Prediction**

### **Computational Intelligence and Neuroscience**

Received 17 October 2023; Accepted 17 October 2023; Published 18 October 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### **References**

- [1] X. You, Y. Liu, M. Zhang, M. Zhang, Y. Yu, and C. Sang, "Modeling on Social Health Literacy Level Prediction," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 7049997, 9 pages, 2021.

## Research Article

# Modeling on Social Health Literacy Level Prediction

**Xuemei You** , **Yongdong Liu**, **Mingming Zhang**, **Man Zhang**, **Yangli Yu**, and **Chengge Sang**

*Business School, Shandong Normal University, Jinan, Shandong 250014, China*

Correspondence should be addressed to Xuemei You; 412095307@qq.com

Received 29 September 2021; Accepted 18 October 2021; Published 7 December 2021

Academic Editor: Daqing Gong

Copyright © 2021 Xuemei You et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Nowadays, the health level of residents has become the focus of people's attention. Under the background of the development of health service from "disease-centered" to "health-centered," it is very important to improve the level of urban health and clarify the factors affecting urban health. Therefore, this paper quantifies the relationship between residents' health literacy level and environment, average life expectancy, infectious disease mortality, and other indicators by selecting appropriate indicators and establishing a mathematical model. Based on the reciprocal linear combination of the collected index data and the corresponding health level value, the prediction model of social health literacy level (SPM) was established, and the qualitative prediction and quantitative analysis of citizens' health literacy level were studied in depth. Based on the SPM model, we can roughly predict the level of health literacy in a region only based on the main variables identified in this paper. The consistency of the experiment shows that the model is effective and robust, and it reveals that environmental factors are the most important factors affecting residents' health literacy level. The actual data show that THE SPM model is a timely and reasonable framework to measure the health literacy level of residents.

## 1. Introduction

Health literacy refers to the ability of individuals to acquire and understand health information and to use it to maintain and promote their own health, including basic knowledge and ideas, healthy lifestyle and behavior, and basic skills. It is an evaluation index that comprehensively reflects the development of national health undertakings [1, 2].

Health literacy is not an isolated concept. Health literacy is related to education [3–7], age [8], race [9], air pollution [10], chronic diseases [11, 12], social medical system [13], etc. However, due to the lack of objective indicators to reflect the overall health literacy level of residents in a certain region or country, the relationship between these indicators and the health literacy level of residents has not been well understood, which is often controversial. Therefore, it is not only of great significance in scientific research, but also of great value in helping people live a healthy life to quantify various indicators and establish a quantitative model to evaluate the health literacy level of residents in a region or country.

With the development of society, people pay more and more attention to the level of residents' health literacy. For the tool to measure health literacy [14], some use experimental methods to evaluate health literacy level [15], some use Delphi survey technology [16], some use HLS-EU-Q questionnaire to measure health literacy level [17, 18], some use mixed multicriteria decision-making (MCDM) method to establish evaluation model [19], and some use dynamic factor model (DFM) [20].

Taking Shenzhen as an example, this paper discusses the relationship between indicators and residents' health literacy level from the perspective of data analysis. By visiting government websites such as Health and Family Planning Commission, Statistical Bureau, Baidu, Souhu, Xinlang, and other search engines and news websites, this paper explores the relationship between health literacy and its multiple variables and makes a quantitative study of their correlation.

In order to explore the influence of many variables on health literacy, we selected 12 factors related to the level of health literacy in combination with 66 articles of health



literacy of Chinese citizens. Then the main variables are obtained by fitting and screening. Using search engine results or data on social media as proxy shows that the main variables do greatly affect the real value of health literacy. The root of this phenomenon lies in three aspects: influencing citizens' basic knowledge and concept, healthy lifestyle and behavior, and basic skills.

Several factors were selected to study the time series of health literacy, and the relationship between the indicators and the level of health literacy was explored. The results showed that PM2.5 (microgram cubic meters), health expenditure accounted for local financial expenditure, infectious disease mortality, and average life expectancy had the highest goodness of fit.

The paper is divided into six parts: the first part is the introduction. The second part includes data source, determination of variables, main variables, and time interval. In the third part, we studied four main variables related to health literacy: PM2.5 (microgram cubic meters), health expenditure, infectious disease mortality, and life expectancy. Then, in the fourth part, we constructed a health literacy level prediction model based on fitting results. The fifth part examines the accuracy and stability of the prediction model. Finally, the conclusions and discussions are presented in Part VI.

## 2. Data

Thirteen sets of data collected from the National Health and Family Planning Commission and the Local Health and Family Planning Commission from 2000 to 2017 were used for research in Shenzhen and the whole country. We will analyze the indicators affecting residents' health prediction level through the trend and fitting of each variable in the past 18 years.

**2.1. Data Sources.** The National Health and Family Planning Commission (NHPC) includes the real values of national health level and the annual changes of various factors over the years. The real health level of local cities can be collected from the local health and family planning committee. Baidu, 360, and other large engines provide all kinds of news reports on health literacy level. At the same time, CCTV, Sohu, and Sina also broadcast in real time. In addition, local TV stations review the course of health literacy in the light of their own characteristics and development trends, in order to continuously improve and look forward to the future.

**2.2. Definition of Variables.** Through a thorough understanding of Article 66 of Chinese Citizens' Health Literacy, combining basic knowledge and concept, healthy lifestyle and behavior, and basic skills, we have established 13 research directions of residents' health literacy  $S(t)$ , maternal mortality  $P(t)$ , the ratio of household registration to non-household registration  $H(t)$ , the number of college graduates  $A(t)$ , the number of hospitals  $N(t)$ , the drinking water standard rate  $W(t)$ , the number of beds  $B(t)$  per 1,000 population, the number of people participating in fitness

activities  $E(t)$ , infant mortality  $I(t)$ , PM2.5  $M(t)$ , health expenditure  $F(t)$ , infectious disease mortality  $G(t)$ , and life expectancy  $L(t)$ . Twelve basic variables except  $S(t)$  were established.

By fitting the basic variables and comparing the trend charts of each factor and health literacy level from 2000 to 2017, four variables with higher fitting degree are obtained, including PM2.5  $M(t)$ , health expenditure accounting for local financial expenditure  $F(t)$ , infectious disease mortality  $G(t)$ , and life expectancy  $L(t)$ .

**2.3. Time Interval  $\Delta t$ .** The true value of health literacy has been calculated since 2000. Every year, the relevant departments will test the level of citizens health literacy nationwide and locally, with an interval of one year. Therefore, we choose one year as our time interval ( $\Delta t$ ), which is consistent with the frequency of the government updating the true value and ranking of health literacy.

## 3. Relevant Variables

**3.1. Variable Screening.** In order to find out the optimal model which can predict residents' health literacy level, we first considered the 12 factors of maternal mortality  $P(t)$ , percentage of household registration and nonhousehold registration  $H(t)$ , number of college graduates  $A(t)$ , number of hospitals  $N(t)$ , drinking water compliance rate  $W(t)$ , number of beds per 1,000 people  $B(t)$ , number of participants in fitness (more than 1000)  $E(t)$ , infant mortality rate  $I(t)$ , PM2.5 content  $M(t)$ , health expenditure accounts for local financial expenditure  $F(t)$ , infectious disease mortality rate  $G(t)$ , and average life span  $L(t)$ .

In this paper, 12 considered variables were fitted with the health literacy level of residents, and the results can be seen in Figure 1. We found that the optimal model was fitted by four variables: PM2.5 content  $M(t)$ , health expenditure accounts for local financial expenditure  $F(t)$ , infectious disease mortality rate  $G(t)$ , and average life span  $L(t)$ .

**3.2. The Health Literacy Level of Residents  $S(t)$ .** The level of residents' health literacy reflects the health literacy status of residents in a country or a region. This article uses the data published by the National Health and Wellness Committee and the Local Health and Wellness Committee as the standard. This paper collected the national residents' health literacy level  $C(t)$  and Shenzhen residents' health literacy level  $S(t)$  between 2000 and 2017, as shown in Figure 2.

Figure 2 shows the development trend of the national residents' health literacy level  $C(t)$  and Shenzhen residents' health literacy level  $S(t)$  from 2000 to 2017. We found that the health literacy level of residents in Shenzhen and the whole country is in a growth trend. In 2016, two solid lines intersect. Figure 2 shows that the national residents' health literacy level  $C(t)$  is higher than the Shenzhen residents' health literacy level  $S(t)$  from 2000 to 2015, and the Shenzhen residents' health literacy level  $S(t)$  is higher than the national residents' health literacy level  $C(t)$  from 2016 to 2017 (see the yellow line section). The intersection point corresponds to

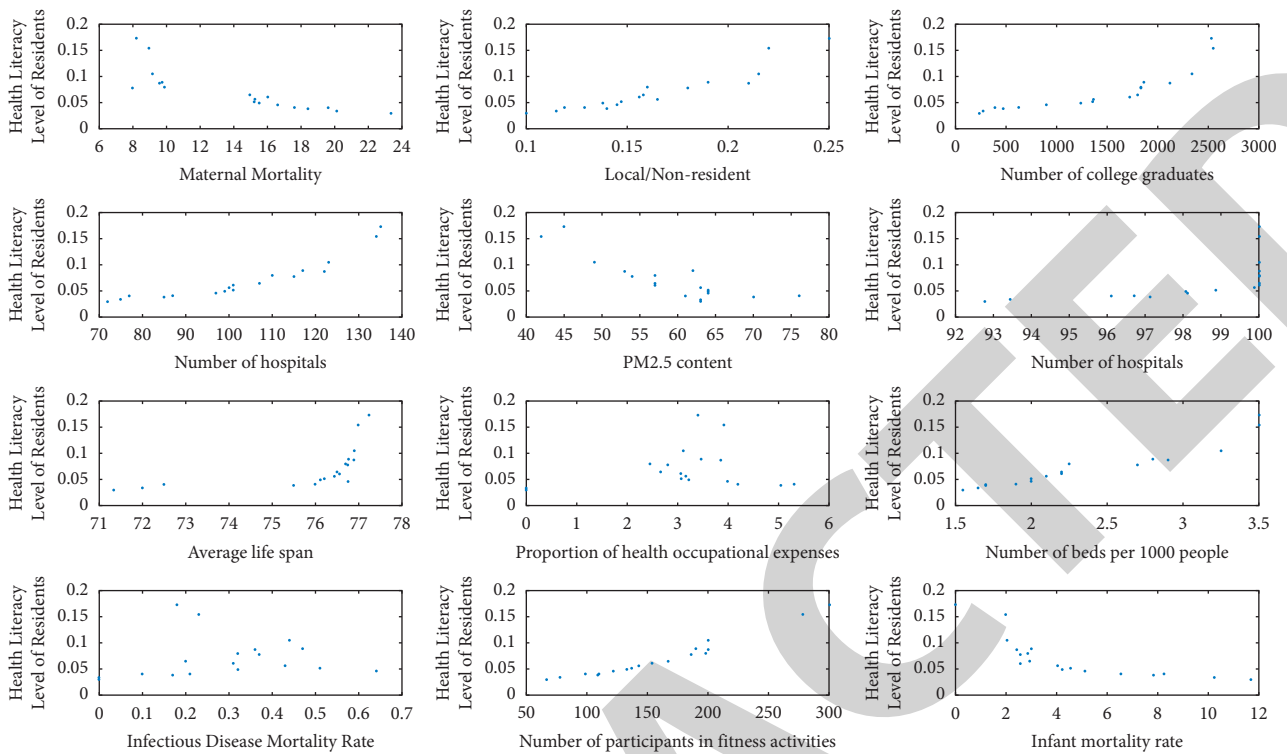


FIGURE 1: Scatter plot of 12 variables and residents' health literacy level.

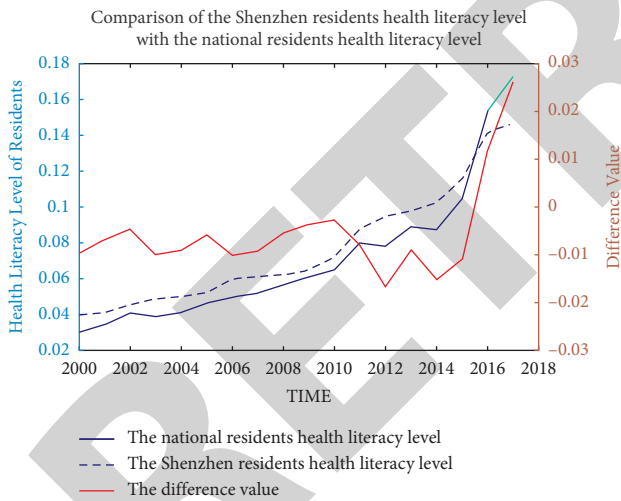


FIGURE 2: The illustration of Shenzhen residents' health literacy level and national residents' health literacy level, where the difference  $D(t)$  between Shenzhen residents' health literacy level and the national residents' health literacy level, national residents' health literacy level  $C(t)$ , and Shenzhen residents' health literacy level  $S(t)$  are shown in dotted red, blue, and red, respectively. The  $x$ -coordinate is time. The left  $y$ -coordinate is the national residents' health literacy level  $C(t)$  and Shenzhen's health literacy level  $S(t)$ . The right  $y$ -coordinate is the difference  $D(t)$  between Shenzhen residents' health literacy level and the national residents' health literacy level.

the Shenzhen Municipal Government's vigorous development of people's livelihood in 2016, which is "to improve the quality of people's livelihood and enhance the level of

people's livelihood security." The document "The implementation opinions of Shenzhen Municipal Government on deepening the reform of medical and health system and building a strong city of health" was issued, along with the reform of medicine and health to establish a higher quality medical and health service system. In 2016, the Shenzhen government also issued "The Shenzhen Solid Waste Pollution Prevention and Control Action Plan" to promote environmental protection and ecological civilization construction in Shenzhen. Therefore, the results show that the residents' health literacy level is closely related to healthcare and ecological environment.

3.3. *The Infectious Disease Mortality Rate  $G(t)$ .* The infectious disease mortality rate  $G(t)$  is an index reflecting the severity of life-threatening diseases, indicating the frequency of death due to a disease in a certain period of time. The incidence of infectious diseases is an indicator of the severity of life-threatening diseases. Influencing factors of reducing mortality rate of infectious diseases in a certain area include improvement of sanitary conditions, improvement of water sources, education (especially women's education), provision of medical services and infrastructure construction. These factors can also obviously affect people's physical and mental health. Therefore, it is reasonable to take the mortality rate of infectious diseases as one of the factors to measure the health literacy level of residents in a certain area. Figure 3 shows the trend of infectious disease mortality and residents' health literacy. As shown in the figure, the lower the mortality rate of infectious diseases, the higher the level of health literacy in the region. When  $G(t) > 0.4$ , the data

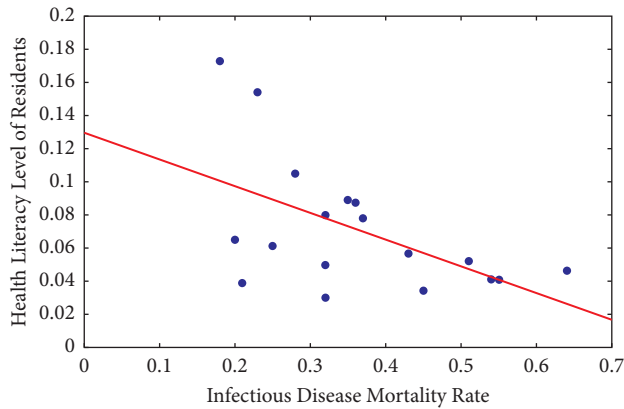


FIGURE 3: The scatter plot of the relationship between the mortality rate of infectious diseases  $G(t)$  and the health literacy level  $S(t)$  of residents from 2000 to 2017, and this graph fits a regression line.

points are more concentrated and fit well with the fitting line. When  $G(t) < 0.4$ , the distribution of data points is scattered and the fitting degree is not as good as before. This may be because with the improvement of medical and health conditions, the mortality rate of infectious diseases is decreasing dramatically and tends to be stable, so the influence of the mortality rate of infectious diseases on the health literacy level of residents is weakening.  $R^2 = 0.333$ . Therefore, we can choose the mortality rate of infectious diseases as an index to measure the health literacy level of residents.

**3.4. The Health Expenditure Accounts for Local Financial Expenditure  $F(t)$ .** Health expenditure refers to the financial allocation for health services by governments at all levels. Health expenditure includes public health service funds and public medical expenses. The proportion of health expenditure in local financial expenditure shows the relationship between public resources consumed by a country (or region) for medical services and those consumed by other public services in a certain period of time. The relationship between health expenditure and residents' health literacy is shown in Figure 4. Figure 4 shows that when  $F(t) > 4$ , health expenditure is closely related to residents' health literacy level, which may be due to the importance of government intervention when residents' health literacy level is relatively low. This also explains the negative correlation between health expenditure and residents' health literacy: when residents' health literacy level is low, people have weak self-awareness, and they need to rely on the government to increase health expenditure, and when residents' health literacy level is high, they can also reduce expenditure. The scatter plot fits the straight line, and its  $R^2$  is 0.202. We choose health expenditure as an indicator of residents' health literacy.

**3.5. The PM2.5 Content  $M(t)$ .** PM2.5 refers to particles smaller than or equal to 2.5 microns in ambient air. PM2.5 can carry a large number of harmful substances through the nasal cavity, directly into the lungs, or even into the blood, so

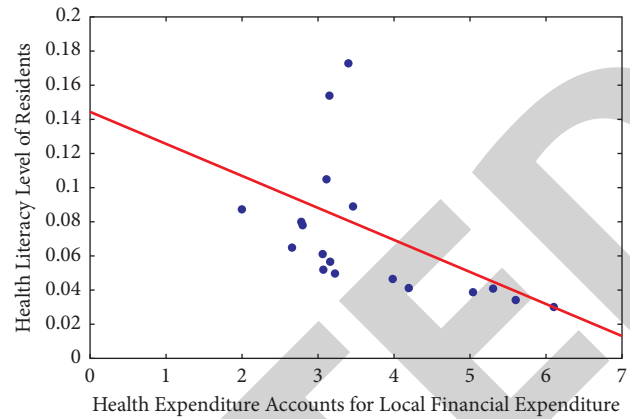


FIGURE 4: A scatter plot of the relationship between health expenditure in local financial expenditure  $F(t)$  and residents' health literacy  $S(t)$  from 2000 to 2017. The straight line in the graph is a fitting line fitted by the scatter plot.

PM2.5 is also known as particulate matter into the lungs, which is closely related to lung cancer, asthma, and other diseases. PM2.5 is the main killer of black lung and haze days and has great harm to human health. The relationship between PM2.5 content and residents' health literacy level is shown in Figure 5. Figure 5 shows that the higher the PM2.5 content, the lower the health literacy level of residents. Therefore, PM2.5 content is negatively correlated with residents' health literacy level. The scatter plot shows that, except for individual points, scatter points are mostly distributed around the fitting line, and  $R^2 = 0.399$ . We regard PM2.5 content as one of the indicators to measure residents' health literacy level.

**3.6. The Average Life  $L(t)$ .** Average life is the average number of years that a person can continue to live in the same period. The life expectancy index comprehensively reflects the level of disease prevention and health services in a country or region. Average life is generally regarded as an important index to measure the quality of life and medical and health level of residents in a country or region and also an important index to evaluate the quality of life and medical level of population in a country or region. The relationship between life expectancy and residents' health literacy is shown in Figure 6. Figure 6 shows that, with the rapid development of regional economy, the remarkable improvement of medical treatment level, and the improvement of people's material living standard, the average life expectancy of population has been growing steadily and rapidly in recent years. But there are several special points that influence this trend. The red dot in Figure 6 corresponds to the residents' health literacy level in 2016 and 2017. In the past two years, the government has fully implemented the project of benefiting the people and actively implemented the strategic plan of building a beautiful China; thus the residents' health literacy level has been significantly improved.

The fitting line  $R^2$  is 0.323 based on scatter plot. We choose life expectancy as one of the indicators to measure the health literacy level of residents in a certain area.

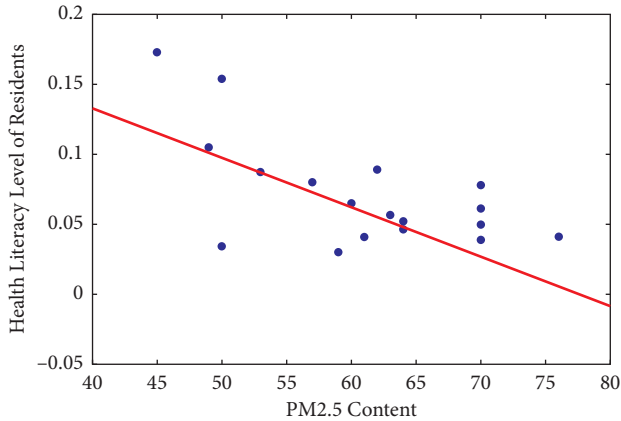


FIGURE 5: A scatter plot of the relationship between PM2.5 content and health literacy level from 2000 to 2017. The straight line in the graph is a fitting line which is fitted by scatter plot.

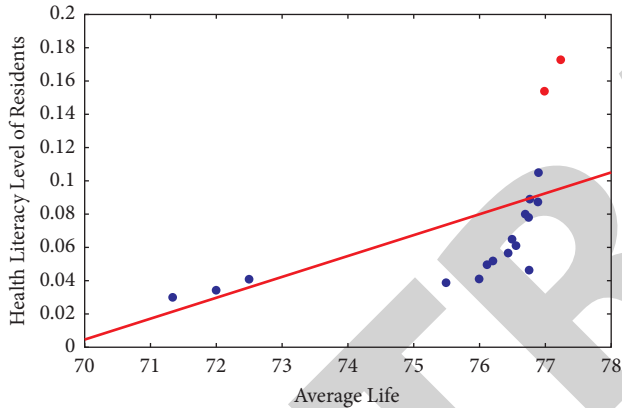


FIGURE 6: A scatter plot of the relationship between life expectancy and health literacy between 2000 and 2017. The straight line in the graph is a fitting line which is fitted by scatter plot.

#### 4. Health Literacy Level Model of Residents

The fourth part shows that the content of PM2.5  $M(t)$ , the proportion of health expenditure in local financial expenditure  $F(t)$ , the mortality rate of infectious diseases  $G(t)$ , and the average life span  $L(t)$  all affect the health literacy level of residents. However, a linear fit  $y(t) = Ax(t) + C$  of each individual indicator measure to the observed residents' health literacy level results in  $R^2 < 0.3$ , except for  $1/M(t)$ , where  $R^2 = 0.361$ . Therefore, there is no single indicator that can adequately measure the health literacy level of residents. This indicates that it is necessary to measure the health literacy level of residents through the combination of indicators (see Appendix for the correlation of variables).

We therefore explored the predictive power of the sum of these variables with multipliers obtained via an ordinary least squares (OLS) fitting process, resulting in  $R^2 = 0.78$ ; the  $p$  values of  $F$ -test for all variables are shown in Table 1. The  $p$  values of variables  $1/G(t)$  and  $1/F(t)$

TABLE 1:  $p$  values and  $\beta$ -coefficients.

	$1/M(t)$	$1/G(t)$	$1/F(t)$	$L(t)$
$p$ value	$\leq 0.001$	0.179	0.423	0.03
$\beta$ -coefficients	0.639	0.183	-0.13	0.621

were greater than 0.05, so the two variables did not pass the test.

We found that the four indexes multiplied and then fitted variables can pass the test and the prediction ability is not much different from the above model. Fitting formula

$$(SPM). S_M(t) = A \frac{F(t)}{M(t)} \frac{L(t)}{G(t)} + C. \quad (1)$$

$R$ -square is 0.640. The  $p$  value of variable  $F$ -test is less than 0.001, which passes the test, so the SPM model is accepted. That is to say, the SPM model is the best model in the models we tested.

According to model 1, we assessed that the impact of that indicator on the health literacy level of residents was greater. Standardized  $\beta$ -coefficient is the corresponding regression coefficient in the regression equation calculated after data standardization [21, 22]. The standardized  $\beta$ -coefficient eliminates the influence of the unit of dependent variable and independent variable [23, 24], and its absolute value can directly reflect the influence degree of independent variable on dependent variable. Table 1 show that the content of PM2.5 is the strongest factor affecting the health literacy level of residents.

#### 5. Experiments

Several experiments were designed to verify the predictive ability of the model SPM:

Firstly, we use the data of the first ten years as training data to get the model coefficients  $A = 0.329$  and  $C = -0.516$ . Then we use this model to predict the health literacy level of residents in the next seven years (Figure 7). We found that this model can accurately predict the health literacy level of residents. Figure 7(a) shows that we compare the predicted value with the real value and find that the scatter points are around the line  $y=x$ , which indicates that the predicted value  $S_M(t)$  is closely related to the real value  $S(t)$ . Figure 7(b) shows the difference between the observed and the real values.  $\Delta S_M(t) = S(t) - S_M(t)$ . The scatter points are distributed around the straight line on the day of  $X=0$  and the fluctuation is small, which indicates that the difference between the predicted value and the real value of the model is acceptable. Figure 7(c) is a comparison between the 95% prediction interval and the true value. We can see from the figure that most of the values are included in the prediction interval. There is a very low deviation in residents' health literacy level because in some years, the official website of the government did not publish its exact value, so there may be permissible error in the value

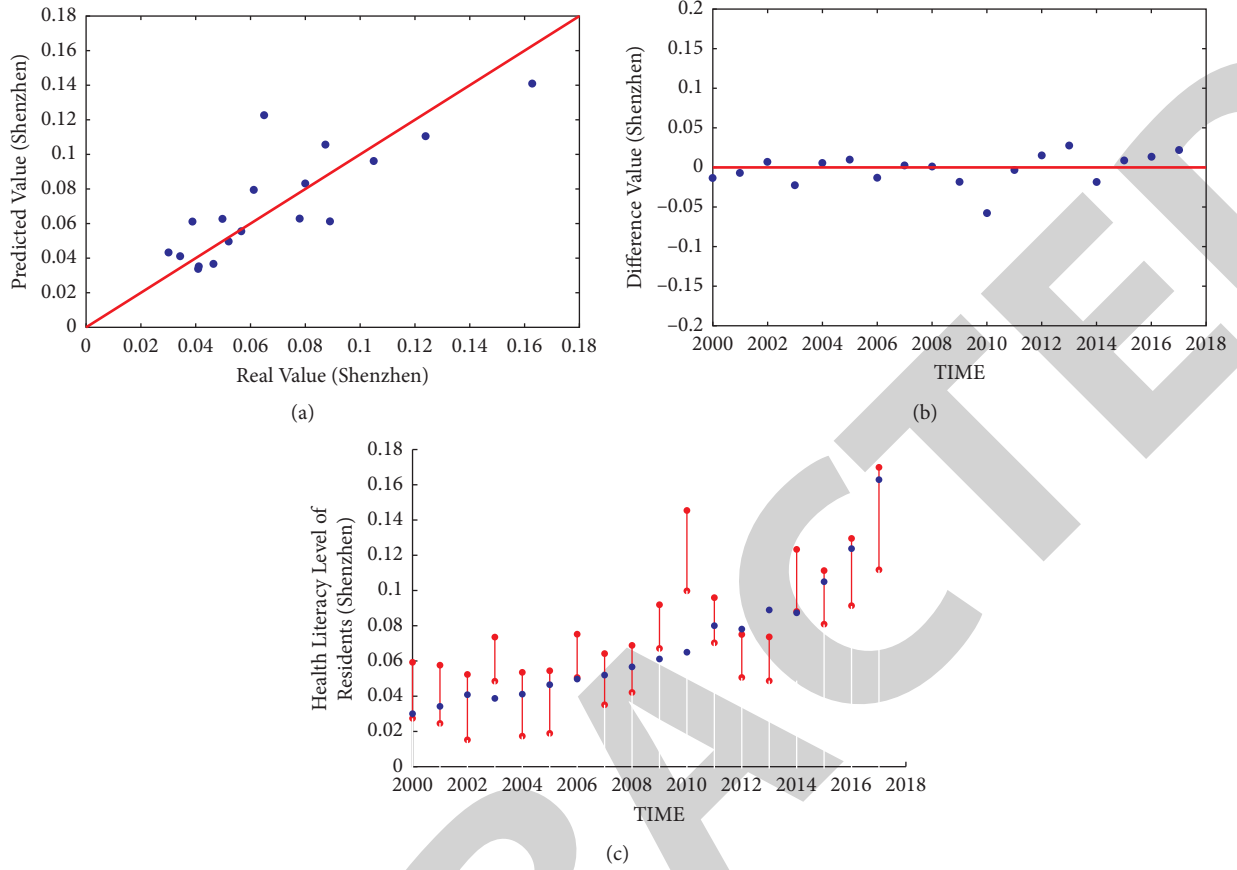


FIGURE 7: Predicting the health literacy level of residents. (a) Based on the data of the previous decade to determine the parameters  $A = 0.329$  and  $C = -0.516$  and to make a scatter plot comparison between the predicted value and the real value. The straight line in the figure is  $y = X$ . (b) is a scatter plot of difference: the difference between the real value and the predicted value is plotted as a scatter plot. Each warehouse (box) in (c) represents the predicted range within a 95% confidence interval.

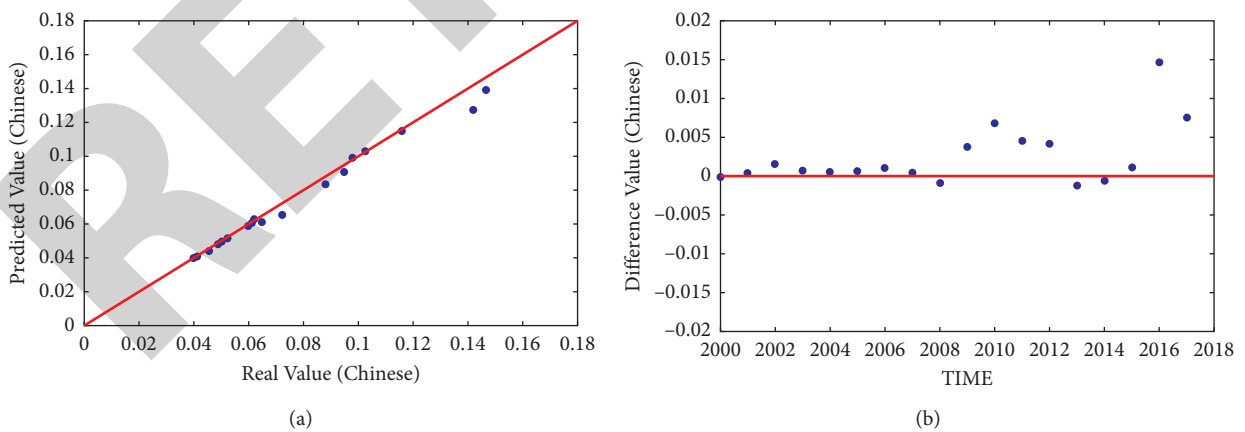


FIGURE 8: Continued.



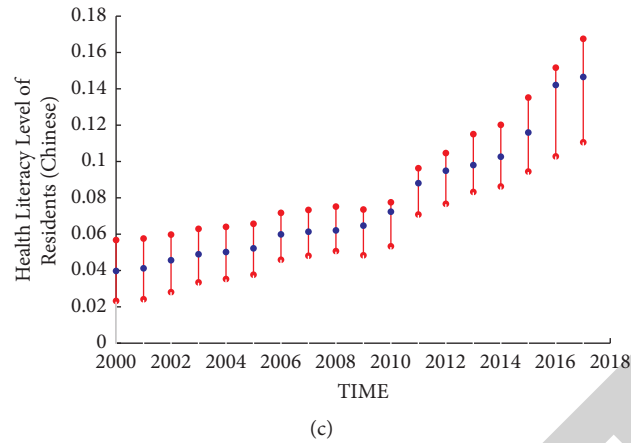


FIGURE 8: According to SPM model, the national residents' health literacy level from 2000 to 2017 is predicted. (a) Scatter plot of predicted and real values. The straight line in the plot is  $x=y$ . (b) Difference scatter plot, in which the points are the difference between the real value and the predicted value. Each warehouse (box) in (c) represents the predicted range within a 95% confidence interval.

TABLE 2: Variable correlation.

		$L(t)$	$1/F(t)$	$1/G(t)$	$1/M(t)$
$L(t)$	Pearson	1	0.721**	0.299	-0.012
	Significance (bilateral)		0.001	0.228	0.961
	$N$	18	18	18	18
$1/F(t)$	Pearson	0.721**	1	0.177	0.124
	Significance (bilateral)	0.001		0.482	0.623
	$N$	18	18	18	18
$1/G(t)$	Pearson	0.299	0.177	1	0.370
	Significance (bilateral)	0.228	0.482		0.130
	$N$	18	18	18	18
$1/M(t)$	Pearson	-0.012	0.124	0.370	1
	Significance (bilateral)	0.961	0.623	0.130	
	$N$	18	18	18	18

\*\*Significant correlation was found at 0.01 level (bilateral).

of residents' health literacy level. Overall, Figure 7 shows that once the PM2.5 content  $M(t)$ , health expenditure accounts for local financial expenditure  $F(t)$ , infectious disease mortality rate  $G(t)$ , and average life span  $L(t)$  are obtained in a certain area, the health literacy level of residents in this area can be predicted. Next, we will carry out further experiments.

Next, we use the national data for further experimental proof. We use SPM model to predict the health literacy level of Chinese residents from 2000 to 2017. We can see from Figure 8(a) that the scatters of predicted and real values are distributed around the  $Y=X$  straight line, which shows that the two values are closely related. Figure 8(b) shows that the difference between the predicted value and the real value fluctuates at  $Y=0$ , and the floating range is small, which further shows that the predicted value of this model has strong predictive ability. Figure 8(c) shows that the model is capable of capturing real values. Therefore, we can say that SPM model has a

good predictive effect on residents' health literacy level.

## 6. Discussion and Conclusions

In recent years, the National Health and Family Planning Commission has incorporated the evaluation index of residents' health literacy into the national health development plan as an evaluation index to comprehensively reflect the development of national health. In this study, we try to quantify the relationship between residents' health literacy and indicators and predict residents' health literacy by indicators.

Because residents' health literacy covers a wide range of areas, it is difficult to obtain it directly. So we selected 12 variables, and after screening, we got four indexes with the highest fitting degree—PM2.5  $M(t)$ , health expenditure accounted for local financial expenditure  $F(t)$ , infectious disease mortality  $G(t)$ , and life expectancy  $L(t)$ . These four indicators can be measured separately. Therefore, this paper establishes a prediction model of residents' health literacy (SPM) based on four indicators. We find that when we input the index data, the



difference between the predicted value and the actual value produced by the model is very small, even neglected. For individual outlier data, we investigated the local information and found that the government website did not publish its exact value, so there may be permissible error in the value of their residents' health literacy level.

Of course, with the rapid development of society, various health problems are constantly emerging. We still have a long way to go in the future on how to predict and control the response and how to find new indicators to measure the level of health literacy in order to continuously improve the model in this paper. At the same time, we hope that the SPM model in this paper can be easily extended to other areas where indicators are easy to measure. For example, health monitoring, land planning, and decision-making [25], asset accumulation and portfolio decisions in the financial sector [26, 27], natural gas demand forecasting [28], food safety, and other indicators are similar to health literacy tests. At the same time, it is hoped that this study can help improve the residents' health literacy level, correctly grasp the basic knowledge and concepts, master healthy lifestyle and behavior, and learn the basic skills of health first aid.

## Appendix

### Variable correlation

Table 2 shows the correlation results of four variables: PM2.5 content  $M(t)$ , proportion of health expenditure in local financial expenditure  $F(t)$ , mortality rate of infectious diseases  $G(t)$ , and average life expectancy  $L(t)$ . The results show that no single index can fully measure the level of health literacy of residents; it is necessary to combine indicators to measure the level of health literacy of residents.

### Data Availability

Thirteen sets of data collected from the National Health and Family Planning Commission and the Local Health and Family Planning Commission from 2000 to 2017 were used for research in Shenzhen and the whole country.

### Conflicts of Interest

The authors declare that they have no conflicts of interest.

### Acknowledgments

This work was supported by the National Natural Science Foundation of China (nos. 71701115, 72171136), the Ministry of Education of Humanities and Social Science Project (no. 21C10445029), the National Social Science Foundation of China (no. 21BGL001), and Shandong Natural Science Foundation (ZR2020MG003).

### References

- [1] D. W. Baker, "The meaning and the measure of health literacy," *Journal of General Internal Medicine*, vol. 21, no. 8, pp. 878–883, 2006.
- [2] D. N. Berkman, T. C. Davis, and L. McCormack, "Health literacy: what is it," *Journal of Health Communication*, vol. 2, no. 15, pp. 9–19, 2010.
- [3] D. Nutbeam, "Health literacy as a public health goal: a challenge for contemporary health education and communication strategies into the 21st century," *Health Promotion International*, vol. 15, no. 3, pp. 259–267, 2000.
- [4] I. S. Kickbusch, "Health literacy: addressing the health and education divide," *Health Promotion International*, vol. 16, no. 3, pp. 289–297, 2001.
- [5] D. Chinn, "Critical health literacy: a review and critical analysis," *Social Science & Medicine*, vol. 73, no. 1, pp. 60–67, 2011.
- [6] I. van der Heide, J. Wang, M. Droomer, P. Spreeuwenberg, J. Rademakers, and E. Uiters, "The relationship between health, education, and health literacy: results from the dutch adult literacy and life skills survey," *Journal of Health Communication*, vol. 18, no. 1, pp. 172–184, 2013.
- [7] T. L. Sentell and H. A. Halpin, "Importance of adult literacy in understanding health disparities," *Journal of General Internal Medicine*, vol. 21, no. 8, pp. 862–866, 2006.
- [8] D. W. Baker, J. A. Gazmararian, J. Sudano, and M. Patterson, "The association between age and health literacy among elderly persons," *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 55, no. 6, pp. S368–S374, 2000.
- [9] N. R. Sperber, H. B. Bosworth, C. J. Coffman et al., "Differences in osteoarthritis self-management support intervention outcomes according to race and health literacy," *Health Education Research*, vol. 28, no. 3, pp. 502–511, 2013.
- [10] H. Kan, S. J. London, G. Chen et al., "Season, sex, age, and education as modifiers of the effects of outdoor air pollution on daily mortality in Shanghai, China: the public health and air pollution in Asia (PAPA) study," *Environmental Health Perspectives*, vol. 116, no. 9, pp. 1183–1188, 2008.
- [11] J. A. Gazmararian, M. V. Williams, J. Peel, and D. W. Baker, "Health literacy and knowledge of chronic disease," *Patient Education and Counseling*, vol. 51, no. 3, pp. 267–275, 2003.
- [12] T. C. Davis, M. V. Williams, E. Marin, R. M. Parker, and J. Glass, "Health literacy and cancer communication," *CA: A Cancer Journal for Clinicians*, vol. 52, no. 3, pp. 134–149, 2002.
- [13] K. Sorensen, S. V. Broucke, J. Fullam et al., "Health literacy and public health: a systematic review and integration of definition of definitions and models," *BMC Public Health*, vol. 12, 2012.
- [14] A. Pleasant, J. McKinney, and R. V. Rikard, "Health literacy measurement: a proposed research agenda," *Journal of Health Communication*, vol. 16, no. 3, pp. 11–21, 2011.
- [15] D. W. Baker, M. V. Williams, R. M. Parker, J. A. Gazmararian, and J. Nurss, "Development of a brief test to measure functional health literacy," *Patient Education and Counseling*, vol. 38, no. 1, pp. 33–42, 1999.
- [16] L. T. Martin, T. Ruder, J. J. Escarce et al., "Developing predictive models of health literacy," *Journal of General Internal Medicine*, vol. 24, no. 11, pp. 1211–1216, 2009.
- [17] K. Sorensen, S. Van den Broucke, J. M. Pelikan et al., "Measuring health literacy in populations: illuminating the design and development process of the European health literacy survey questionnaire (HLS-EU-Q)," *BMC Public Health*, vol. 13, no. 1, 2013.
- [18] T. V. Duong, A. Aringazina, G. Baisunova et al., "Measuring health literacy in Asia: validation of the HLS-EU-Q47 survey tool in six Asian countries," *Journal of Epidemiology*, vol. 27, no. 2, pp. 80–86, 2016.

## Research Article

# A Mixed Decision Strategy for Freight and Passenger Transportation in Metro Systems

Yutao Ye, Junhua Guo , and Lixin Yan

*School of Transportation and Logistics, East China Jiaotong University, Nanchang, Jiangxi Province 330013, China*

Correspondence should be addressed to Junhua Guo; gjhtougao@163.com

Received 21 October 2021; Revised 3 November 2021; Accepted 17 November 2021; Published 1 December 2021

Academic Editor: Daqing Gong

Copyright © 2021 Yutao Ye et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper proposes a mixed decision strategy for freight and passenger transportation in metro systems during off-peak hours (MTS-OPH). The definition of the mixed decision strategy is proposed, and fixed and flexible loading modes are considered for different passenger flow volumes. A mathematical model of the MTS-OPH is proposed and solved using an improved variable neighborhood search algorithm. Case studies demonstrate the performance and applicability of the proposed model and algorithm, and the MTS-OPH is discussed for different delivery distances, passenger flows, and metro network types. The proposed strategy is suitable for long-distance delivery, and the proposed model framework can be applied to different types of metro networks with different levels of complexity. The mixed decision strategy provides a decision support tool for metro and freight companies and can propose corresponding solutions according to different passenger flows.

## 1. Introduction

With the continuous development of e-commerce and home deliveries, urban freight transport has emerged as a key link in urban economic and social development [1, 2]. Road transportation is the primary mode for urban freight transportation [3], and 85–90% of freight is transported by road in France [4]. However, the freight transportation by vehicles poses a series of problems pertaining to urban traffic congestion, greenhouse gas emissions, and noise. Urban freight accounts for 10% of total transportation, but it accounts for 40% of urban pollutant emissions [5]. Therefore, optimizing the transportation structure, strengthening the cooperation of different transportation modes, promoting the organic connection between intercity trunk transportation and city-end distribution [6], and encouraging the development of intensive distribution models are of great significance for creating a green and efficient logistics system.

As mentioned above, the freight transportation needs to change towards more efficient and sustainable transportation systems to cope with increasing demand for freight transportation in urban areas. The metro has the advantages of high efficiency, large capacity, and sustainability, but most

metro networks suffer from insufficient utilization of metro trains due to the low passenger flow during off-peak hours [7]. Therefore, introducing goods into the metro network during off-peak hours is a very potential way of freight [8]. The present mixed transport strategy for freight and passenger transportation in metro systems is usually by subjective experience [9], lacking reasonable theoretical framework and mathematical formulation. Furthermore, the entire metro network should be considered to deal with the freight, rather than a single metro line [10]. Therefore, a general theoretical framework and model for constructing the mixed passenger and freight transportation strategy on the metro network during off-peak hours are indispensable.

The purpose of this study is to combine the existing metro network with the first- and last-mile delivery services operated by logistics companies and propose a feasible passenger and freight flow mixed transport strategy for metros during off-peak hours (MTS-OPH). The paper considers fixed and flexible loading modes under different passenger flows and analyzes different cargo characteristics and delivery time requirements. A model for quantitatively evaluating the mixed transportation strategy is established, and an improved variable neighborhood search (VNS) algorithm is designed. They are then applied to the cargo

transportation of Ningbo and Beijing metro networks, respectively, and the mixed transportation strategy under different delivery distances, passenger flows, and metro network types is analyzed. The model and method can also be applied to the metro network of other rail transit cities.

The remainder of this paper is organized as follows: Section 2 reviews related research on metro mixed transportation. Section 3 formulates a nonlinear programming model of the MTS-OPH, and then, an improved VNS algorithm is designed to solve the proposed model in Section 4. Subsequently, two case studies are implemented to verify the proposed model in Section 5. In Section 6, the conclusions of the study and scope for future research are presented.

## 2. Literature Review

*2.1. Urban Freight and Passenger Transportation.* Nash [11] first proposed the use of urban public transportation for freight transportation. Later, based on the concept of sharing, Trentini et al. [12] proposed to introduce urban freight transportation into passenger transportation to achieve the purpose of sharing transportation resources and transportation infrastructure. On this basis, the flow of urban freight and passenger transport was quantified, and a new urban transport system for passenger and freight was constructed and implemented in La Rochelle, France [13]. Gonzalez-Feliu [14] used a socioeconomic cost-benefit analysis to assess the applicability of tram freight transport in the Paris region. The study showed the potential of the tram freight transport mode. Fatnassi et al. [15] integrated personal rapid transportation and freight rapid transportation modes and used electric vehicles to achieve mixed passenger and freight transportation on automatic rails. Masson et al. [16] proposed a hybrid freight method based on the integration of passenger and freight systems to solve mixed urban traffic problems. This method used buses to transport goods from the central distribution center to the transfer point and then used a tricycle to transport the goods from the transfer point to the destination.

There are some successful cases of urban passenger and freight transportation. In Dresden, the tram line from the Volkswagen warehouse to the city center can transport 300,000 tons of goods a year, 10 times a day, which significantly reduces carbon dioxide emissions [17]. In Paris, the commuter line D transports household goods, leisure products, and other goods from the MONOPRIX warehouse (Combs-la-Ville and Lieusaint) on the outskirts of Paris to the Bercy station [18]. These goods are then transported to stores through trucks that use natural gas vehicle fuel to satisfy the emission reduction principles along the logistics line. In New York, subway waste is collected through stations using modified metro trains [19]. However, most of the existing research on mixed transportation modes is based on road and railway transportation. These above studies can be applied into the mixed freight and passenger transportation in the metro system. Since the operation mode, transportation efficiency, and transportation timeliness of metro are different from railways, the mixed

transportation of freight and passenger transportation in the metro systems needs further research.

*2.2. Urban Freight and Passenger Transportation in Metro Systems.* The feasibility and application prospects of urban metro freight have been studied in a few research papers. Rijsenbrij and Pielage [20] discussed the feasibility of using the metro for mixed passenger and freight transportation, which gained the attention of national and international scholars on metro freight transportation. Kikuta et al. [9] conducted a test study on the combination of the public metro service and conventional truck operation to prove the feasibility of this mode of transportation. For the metro distribution service, Motraghi and Marinov [21] collaborated with the Newcastle metro network to theoretically analyze urban freight transportation. However, further research is required before actual implementation of their findings. The above research made some explorations on the mixed transportation of freight and passengers in metro systems but did not establish an implementation ability and systematization theory framework. Therefore, a general mixed transportation theoretical framework in the metro systems should be designed to realize the freight and passenger mixed transportation strategy.

In addition, a few studies focused on metro freight. Brice et al. [22] reported a baggage transfer system to facilitate passengers transporting luggage from Newcastle city center to Newcastle International Airport via the metro. They proved the feasibility of the new baggage transfer solution and reported that the corresponding cost is higher than that of the existing service. Ghilas et al. [23] found that integrated freight and scheduled line services can potentially reduce the operating costs of logistics service providers and the public transportation sector can obtain additional revenue. However, the existing mixed transportation models are established based on a single metro line. Considering the different directions of goods circulation, existing models cannot accurately describe the mixed transport strategy in metro systems. Therefore, a mixed transportation model on the metro network should be established to realize the circulation of goods.

*2.3. Optimization Method of Urban Freight and Passenger Transportation.* Location and route selection have long been the significant issues in the freight and passenger transportation domain. Fatnassi et al. [15] proposed two mathematical formulas to solve the vehicle route between stations and used dynamic optimization methods and developed algorithms to solve the shared transportation of goods and passengers. Zhao et al. [24] proposed a segmentation method of urban metro network, using complex network theory and the TOPSIS model to determine the candidate metro distribution hubs for the location model. Dong et al. [25] analyzed the characteristics of underground cargo capacity and established a mixed integer programming model to select the location of metro distribution hubs. However, the limitation of the frequency of goods transfer in the metro network is barely considered in these studies.



Likewise, time window constraint is a critical issue along with the rapid increase in freight demand during recent years. Behiri et al. [8] studied an environment-friendly urban freight transportation alternative using a passenger railway network and proposed a heuristic based on dispatching rules and a single-train-based decomposition heuristic to solve the Freight-Rail-Transport-Scheduling Problem. Yang et al. [26] studied the vehicle routing problem with mixed backhauls and time windows for city logistics, and the time-dependent pickups and deliveries can be depicted by extending the state dimensions. However, freight transportation in the passenger transportation system focuses on the study of the cargo time window, ignoring the impact on the normal operation time window of the passenger transportation system.

In addition, the development of solving algorithms has attracted the attention of researchers and practitioners. Bräysy [27] introduced the internal design of variable neighborhood descent and variable neighborhood search algorithms in detail, analyzed the vehicle routing problem with time windows problem, and pointed out that the variable neighborhood search algorithm is one of the most effective methods to solve the vehicle routing problem with time windows problem. de Armas and Melián-Batista [28] studied a dynamic rich vehicle routing problem with time windows and proposed a meta-heuristic algorithm based on variable neighborhood search to solve this problem. However, the stability and reliability of the calculation results need to be strengthened. Therefore, we design two types of neighborhood structures to obtain high-quality solutions.

**2.4. Contribution.** The main contributions of this study are summarized as follows:

- (i) We propose a general theoretical framework of MTS-OPH. This framework includes new concept and transportation standards of MTS-OPH. This definition clarifies the process and applicable time of mixed transportation. Relative to previous study on the types of mixed transportation goods [13, 29], MTS-OPH further explored the types of goods suitable for mixed transportation on the basis of transporting small goods. The mixed transportation standard of separate loading of passengers and cargo and priority transportation of passenger flow is established. Furthermore, fixed and flexible cargo loading modes are proposed according to different off-peak passenger flows.
- (ii) We formulate a nonlinear programming model for the MTS-OPH. Based on the research of Fatnassi et al. [15] and Zhao et al. [30], our model improves the mixed transportation strategy research to a mixed transportation strategy research under load rates of passenger flow during off-peak hours. To the best of our knowledge, this is the first time that the problem of mixed passenger and freight transportation under different passenger flows has been addressed in the context of metro transportation during off-peak

hours. Moreover, an improved VNS algorithm is designed to solve the model, which provides a decision support tool for logistics companies.

- (iii) We presented the real case study of the Ningbo and Beijing metro network to verify the practicality and efficiency of the proposed model. The applicability of the model is discussed from the distribution distance, off-peak passenger flow, and metro network type. Our results show that the proposed model can be applied to different metro networks. Furthermore, the mixed transportation mode in the metro systems has the advantages of high speed, high punctuality, low economic investment, low environmental impact, and low energy consumption [31–33].

### 3. Mathematical Formulation

This section provides a detailed description of the MTS-OPH, analyzes the types of goods suitable for mixed transportation, and provides different mixed transportation methods based on different load rates of passenger flow during off-peak hours. Finally, a nonlinear programming model of a mixed transportation strategy is constructed.

**3.1. Description of MTS-OPH.** MTS-OPH integrates the existing urban metro network with the first- and last-mile delivery services operated by logistics firms. The key aspect of this strategy is to transport goods and passengers together without affecting the metro passenger flow. Thus, the transportation standard of separate loading of passengers and cargo and priority transportation of passenger flow should be considered. First, metro train carriages are divided into passenger carriages and freight carriages, similar to the passenger carriages and freight carriages set in the Shenzhen metro trains in China [34]. Different types of carriages are set with boundary lines and dedicated passages to divide the passenger and freight flows, as shown in Figure 1. To minimize the impact of cargo flow on passengers, we introduce relevant constraints on the types of cargo used for mixed transportation. For ease of operation, all goods are placed in a same-size freight box as freight parcels. Concurrently, small cargo packages should be selected for metro mixed transportation. A similar conclusion has been reported in related studies on bus freight [13] and metro freight [9, 29]. Thus, goods in the freight parcel should be small goods, such as documents, books, clothing, and small mechanical parts, rather than fresh goods that need to be frozen and kept fresh, which are not affected by time and the environment, in order to maintain the metro carriage environment tidy.

Second, to satisfy the passenger flow priority transportation standard without changing the headway and stop time of metro trains during off-peak hours, this study proposes different mixed transportation methods based on different load rates of passenger flow during off-peak hours. When the number of passengers and freight demand is greater than the metro capacity, each train adopts a mixed passenger and cargo transportation mode during off-peak

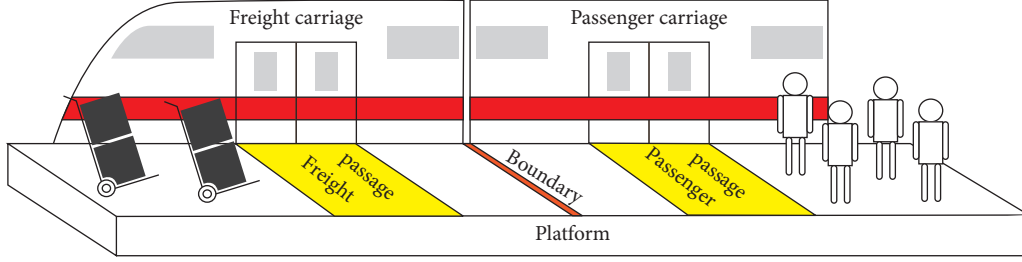


FIGURE 1: Processing of the second stage.

hours, as shown in Method 1 (i.e., fixed loading mode) in Figure 2. When the number of passengers and freight demand is less than the metro capacity, one of every two trains is selected to employ the mixed transportation mode while the other train is used for passenger transportation, as shown in Method 2 (i.e., flexibility loading mode) in Figure 2. In Figure 2,  $t_{\text{interval}}$  represents the headway of metro trains during off-peak hours.

In summary, MTS-OPH is a transportation strategy based on the aforementioned mixed transportation standard, which satisfies the optimal distribution cost of the cargo flow under the condition of determining the origin and terminal stations of the goods. The transportation strategy is to use the metro passenger transportation network, which is composed of cargo distribution centers, metro stations, terminal cargo stations, metro trains, and freight vehicles, as shown in Figure 3.

In Figure 3, the operation process of MTS-OPH is divided into five stages. The first and fifth stages are vehicle delivery, second and fourth stages are manual transshipments, which transfer cargo from the vehicle to metro carriage, and third stage is metro delivery. The transfer of goods between each stage is completed by trolley.

In the first stage, the cargo is packed in freight parcels, placed in freight boxes, loaded onto transportation vehicles, and delivered to the departure metro station via freight vehicles. Then, the freight boxes are unloaded from the vehicle and loaded on the trolley. In the second stage, the freight boxes are transported to the metro platform via trolleys and transported to the metro freight carriage when the train enters the platform. In the third stage, freight boxes are transported on the metro network (transfers are performed via trolleys). Here, the train operation mode is from the first station to the last station regardless of the train service route. The freight boxes enter the fourth stage after being transported via the metro train. The freight boxes are unloaded from the metro freight carriage and transported via trolleys to the arrival metro station. In the last stage, the freight boxes are loaded onto the freight vehicle and transported to the corresponding terminal cargo station. All the aforementioned stages constitute the MTS-OPH operation process, while the third stage of the operation period in MTS-OPH is the off-peak hours of metro operation.

**3.2. Assumptions and Notations.** The MTS-OPH model can be described as the process of delivering multiple freight stations from a cargo distribution center with the optimal delivery cost

as the goal. The number of terminal cargo stations is determined, but the freight stations are located in different geographical locations. All transportation paths are based on the actual shortest distance of the road network or metro network. Thus, certain assumptions have been considered as follows:

- (i) The freight is placed in a standardized unit parcel to measure the arrival demand of the terminal cargo stations.
- (ii) There is no storage function at the metro exit station, and freight vehicles are arranged for delivery immediately after the cargo arrives at the metro exit station.
- (iii) The freight vehicles used for transportation are of the same model with the same fuel consumption and load capacity.
- (iv) The freight vehicles run at a uniform velocity without consideration of the road conditions. Furthermore, after the freight is transported, the vehicles are not required to return to the cargo distribution center or arrival metro stations.
- (v) Up to two freight transfers occur in the urban metro network.

Table 1 summarizes the notations used throughout the paper.

### 3.3. System Constraints

#### 3.3.1. Number of Freight Carriages Constraints

$$E = D - \left\lceil \frac{\max N}{\text{cap}_M} \right\rceil, m = 1, \dots, M. \quad (1)$$

Equation (1) determines the number of freight carriages ( $x$ ) represents the smallest integer greater than  $x$ ).

#### 3.3.2. Freight Loading Mode Constraints

$$r = \begin{cases} 1, & \frac{Q_{i'}}{E \times \text{cap}_M} > 1, \\ 2, & \frac{Q_{i'}}{E \times \text{cap}_M} \leq 1, \end{cases} \quad i' = 1, \dots, I'. \quad (2)$$

The freight loading mode is selected by Equation (2). Among them,  $r = 1$  is a fixed loading mode; that is, each train uses a mixed passenger and freight transportation

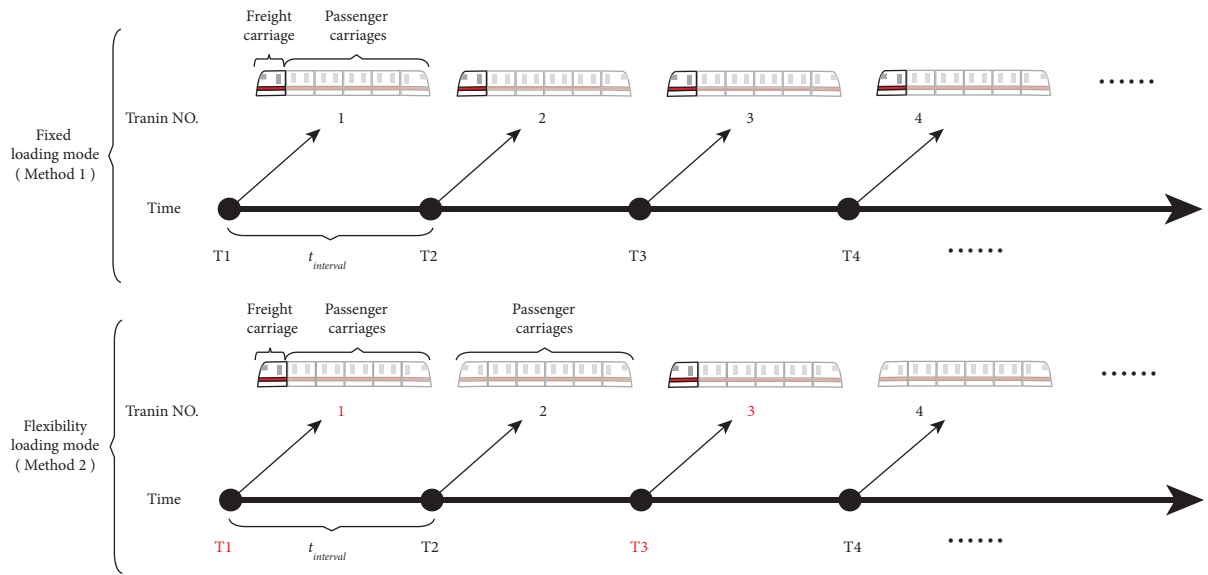


FIGURE 2: Two transportation methods of MTS-OPH.

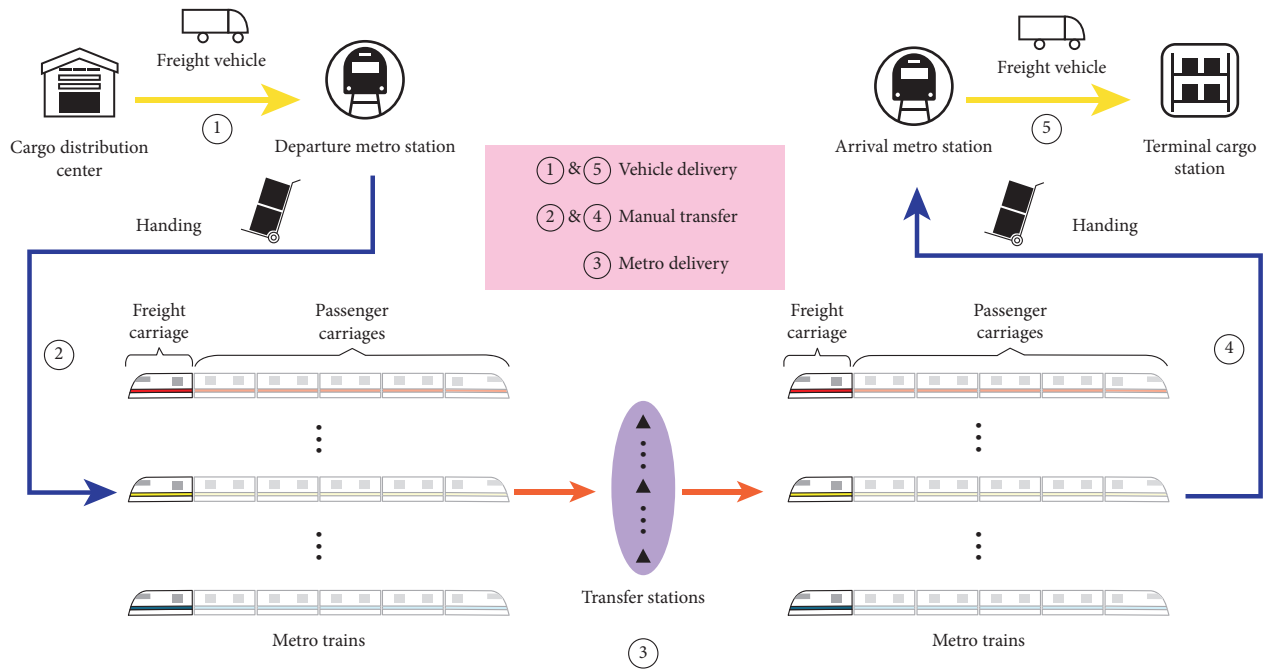


FIGURE 3: Operation process of MTS-OPH.

TABLE 1: Notations.

Index sets	$J$	Set of departure metro stations, $j \in J$
	$J'$	Set of arrival metro stations, $j' \in J'$
	$I'$	Set of terminal cargo stations, $i' \in I'$
	$V$	Set of freight vehicles, $v \in V$
	$M$	Set of metro, $m \in M$
	$N$	Set of maximum cross-section passenger flow collection during off-peak hours, $N_m \in N$ , where $N_m$ is the maximum cross-section passenger flow of metro $m$ during off-peak hours



TABLE 1: Continued.

Parameters	$i$	Cargo distribution center ( $i = 1$ )
	$n$	Index of freight vehicles ( $n \in V$ )
	$Q_{i'}$	Number of cargos received at terminal cargo station $i'$
	$C_v$	Fixed operating cost for one vehicle
	$C_{ij}$	Cost of per unit cargo transportation by vehicles from $i$ to $j$
	$C_O$	Cost of per unit cargo transfer after one transfer
	$t_O$	Time taken to complete a transfer
	$D$	Number of train carriages
	$E$	Number of freight carriages
	$\text{cap}_V$	Capacity of transport vehicles
	$\text{cap}_M$	Freight capacity of one metro carriage
	$O_{jj'}$	Number of transfers on metro transport path ( $j, j'$ )
	$s_j^m$	On the metro line where station $j$ is located, the departure time of the first station of train $m$
	$t_j$	Cargo loading time at station $j$
	$t_{ij}$	Time taken to complete path ( $i, j$ )
	$[\text{ET}_{i'}, \text{LT}_{i'}]$	Receiving time window of the terminal cargo station $i'$
	$r$	1 represents Method 1 of MTS-OPH, 2 represents Method 2 of MTS-OPH
$\alpha$	Delay cost of unit cargo	
$\lambda$	0, if the delivery time is less than the latest delivery time, $\alpha$ , otherwise	
$Z$	A large positive number	
Decision variables	$x_{ij}^v$	1, if the transport path ( $i, j$ ) is served by the vehicle $v$ , 0, otherwise
	$y_{jj'}^{i'm}$	1, if the cargo from the terminal cargo station $i'$ is delivered by the metro train $m$ from the departure station $j$ to the arrival station $j'$ , 0, otherwise
	$\text{SET}_{i'}$	Time taken to start delivery to the terminal cargo station $i'$
	$\text{SLT}_{i'}$	Time taken to complete delivery to the terminal cargo station $i'$
	$K$	Quantity of transport vehicles ( $K \in Z$ )

mode during off-peak hours;  $r = 2$  is a flexible loading mode; that is, one of every two trains chooses to use a mixed transportation mode and the other for passenger transportation.

**3.3.3. Vehicle Line Capacity Constraints.** Vehicles in the “first mile and last-mile” distribution network should meet vehicle capacity constraints, vehicle number constraints, and line capacity constraints.

$$\sum_{v=1}^n x_{ij}^v = 1, \quad i = 1, j = 1, \dots, J, \quad (3)$$

$$\sum_{v=n+1}^V \sum_{j'=1}^{J'} x_{j'i'}^v = 1, \quad i' = 1, \dots, I', \quad (4)$$

$$Q_{i'} \leq \sum_{v=n+1}^V \sum_{j'=1}^{J'} x_{j'i'}^v \text{cap}_V, \quad i' = 1, \dots, I', \quad (5)$$

$$\sum_{j=1}^J \sum_{v=1}^n x_{ij}^v + \sum_{j'=1}^{J'} \sum_{i'=1}^{I'} \sum_{v=n+1}^V x_{j'i'}^v = K, \quad i = 1. \quad (6)$$

Equation (3) assigns a unique route for each vehicle from the cargo distribution center to the departure metro station. Equation (4) imposes the restriction that each vehicle can have one unique route from one arrival metro station to one terminal cargo station in the third stage. Equation (5) constrains the capacity on the route. The total number of

vehicles in the complete distribution process should be restrained, as shown in Equation (6).

**3.3.4. Metro Line Capacity Constraints.** The goods in the metro network by train should meet the train capacity constraint and the number of transfer constraints.

$$Q_{i'} \leq E \times \text{cap}_M \times \sum_{m=1}^M \sum_{j=1}^J \sum_{j'=1}^{J'} y_{jj'}^{i'm}, \quad i' = 1, \dots, I', \quad (7)$$

$$\sum_{j=1}^J O_{jj'} \leq 2, \quad j' = 1, \dots, J'. \quad (8)$$

Equation (7) expresses the capacity limit of the metro. Equation (8) stipulates the number of transfers in the metro network.

**3.3.5. Delivery Time Constraint.** The mixed transportation is based on the off-peak hours of urban rail transit, and the departure time of mixed transportation trains shall not be earlier than the start time of off-peak hours.

$$\sum_{s=1}^{j-1} t_s + s_j^m \leq t_j, \quad j = 1, \dots, J, m = 1, \dots, M, \quad (9)$$

$$t_j - t_{ij} - (1 - x_{ij}^v)Z \leq \text{SET}_{i'}, \quad i = 1, j = 1, \dots, J, v = 1, \dots, n, i' = 1, \dots, I', \quad (10)$$

$$ET_{i'} \leq SET_{i'} \leq LT_{i'}, \quad i' = 1, \dots, I'. \quad (11)$$

Equation (9) determines the earliest time when the freight to cargo station is loaded into the train at a station. In this equation,  $t_s$  represents the running time between  $S$  and  $S-1$  when the train is at station  $S$ , and when  $s=1$ ,  $t_s=0$ . Equation (10) ensures the time to start delivery for the terminal cargo station. Equation (11) is the time window constraint of the terminal cargo station.

The delivery time of the train in the whole mixed transportation process meets the sum of the delivery time of the ‘‘first mile and last-mile’’ and the metro delivery time.

$$\begin{aligned} SLT_{i'} &= SET_{i'} + t_{ij} + (t_{jj'} + r \times t_{\text{interval}}) y_{jj'}^{i'm} + O_{jj'} t_O \\ &\quad + t_{j'i'}, \quad j = 1, \dots, J, \\ j' &= 1, \dots, J', i' = 1, \dots, I', m = 1, \dots, M. \end{aligned} \quad (12)$$

Equation (12) addresses the relationship between the start time and end time of delivery from the cargo distribution center to the terminal cargo station.

The delay cost coefficient is affected by the delivery time window of the goods. Failure to complete the delivery within the time window needs to calculate the delay cost based on the delay cost coefficient.

$$\lambda_{i'} = \begin{cases} 0, & SLT_{i'} \leq LT_{i'}, \\ \alpha, & SLT_{i'} \geq LT_{i'}, \end{cases} \quad i' = 1, \dots, I'. \quad (13)$$

Equation (13) addresses the delay cost due to the failure to complete the delivery within the time window; the delay cost is related to the time of completion of the delivery.

**3.3.6. Decision Variable Constraints.** The relevant decision variables are constrained as follows. Equation (14) expresses the constraint on the initial number of vehicles.

$$K \geq 2, \quad (14)$$

$$SET_{i'}, SLT_{i'} \geq 0, \quad i' = 1, \dots, I', \quad (15)$$

$$x_{ij}^v \in \{0, 1\}, \quad i = 1, \dots, I, j = 1, \dots, J, v = 1, \dots, n, \quad (16)$$

$$\begin{aligned} x_{j'i'}^v \in \{0, 1\}, \quad j' = 1, \dots, J', i' = 1, \dots, J', \\ v = n+1, \dots, V, \end{aligned} \quad (17)$$

$$\begin{aligned} y_{jj'}^{i'm} \in \{0, 1\}, \quad j = 1, \dots, J, j' = 1, \dots, J', \\ i' = 1, \dots, I', m = 1, \dots, M. \end{aligned} \quad (18)$$

**3.4. Composition of the Objective Function.** The objective function of MTS-OPH is composed of vehicle transportation cost, transfer cost, and delay cost. Each cost parameter is described in the following.

**3.4.1. Vehicle Transportation Cost.** The transportation cost of vehicles is mainly composed of the transportation

distance, number of vehicles, number of freight parcels, and fixed operating cost of vehicles. Therefore, the transportation cost  $f_{vij}$  in route  $(i, j)$  is as shown in

$$f_{vij} = Q_{i'} y_{jj'}^{i'm} x_{ij}^v C_{ij}. \quad (19)$$

To express  $f_{vij}$  clearly, Equation (19) is simplified as

$$Q_{ij} = Q_{i'} y_{jj'}^{i'm} x_{ij}^v, \quad (20)$$

$$f_{vij} = C_{ij} Q_{ij}. \quad (21)$$

Therefore, the vehicle transportation cost  $f_{vj'i'}$  in the route  $(j', i')$  can be obtained similarly.

$$f_{vj'i'} = C_{j'i'} Q_{j'i'}. \quad (22)$$

In summary, Equation (23) represents the vehicle transportation cost.

$$f_v = \sum_{i=1}^J \sum_{j=1}^n \sum_{v=1}^V C_{ij} Q_{ij} x_{ij}^v + \sum_{j=1}^J \sum_{i'=1}^{I'} \sum_{v=n+1}^V C_{j'i'} Q_{j'i'} x_{j'i'}^v + KC_v. \quad (23)$$

**3.4.2. Transfer Cost.** The number of freight parcels and frequency of transfer constitute the transit cost. The number of transfers is mainly determined by the number of times the cargo is loaded and unloaded; the completion of one loading and unloading of cargo is regarded as one transfer. Equation (24) shows the calculation method for the node transfer cost  $f_n$ .

$$f_n = 2C_O \left( \sum_{i'=1}^{I'} Q_{i'} + \sum_{j=1}^J \sum_{j'=1}^{J'} \sum_{i'=1}^{I'} \sum_{m=1}^M O_{jj'} y_{jj'}^{i'm} Q_{i'} \right). \quad (24)$$

**3.4.3. Delay Cost.** The delay cost is caused by the freight arriving at the terminal cargo station in an unexpected time window. The delivery time window of MTS-OPH is based on the off-peak start time, so the delay cost will only be calculated if the delivery time window is exceeded. This is expressed with  $f_d$ , as in

$$f_d = \sum_{i'=1}^{I'} \lambda_{i'} Q_{i'}. \quad (25)$$

Because only part of the metro carriages is used for freight, the weight of the cargo is much smaller than that of the metro. Therefore, the delivery cost of MTS-OPH is not affected by metro fixed operating cost and metro transportation distance.

In summary, the delivery cost  $f_C$  is the sum of the vehicle transportation cost  $f_v$ , node transit cost  $f_n$ , and delay cost  $f_d$ , i.e.,  $f_C = f_v + f_n + f_d$ . Hence, the mathematical formulation of the MTS-OPH model is built as follows:

$$\min Z = f_C. \quad (26)$$

Equation (26) denotes the smallest delivery cost, and the calculation process is, respectively, performed using Equations (19)–(25).

In the MTS-OPH model, the pickup and delivery constraints of the time window and transshipment value are considered [8, 35, 36] and the mode selection constraints are added according to the two loading modes of the proposed mixed transport strategy. This model is developed based on the generalized assignment problem, which belongs to a NP-hard problem [37–39] and is usually solved by heuristic algorithms.

## 4. Solution Approaches

The heuristic algorithm for solving the MTS-OPH is based on the VNS [40, 41]. VNS provides a flexible framework for constructing heuristics for approximately solving combinatorial optimization problems and nonlinear optimization problems. The main idea is to dynamically change the neighborhood structure set during the search process to expand the search range and obtain local optimal solutions. Using such a variable neighborhood strategy, it is possible to move away from the optimum and finally reach convergence after multiple iterations. In the VNS algorithm used in this study, the objective function (indicated by  $Z_c$ ) in Equation (8) is used as an evaluation index for evaluating the quality of the generated solution, as shown in the following:

$$f(c) = Z_c. \quad (27)$$

In this study, the VNS algorithm includes the following three parts: initial solution, shaking process, and variable neighborhood descent (VND) process. We use  $c$ ,  $c'$ , and  $c''$  to denote the solutions generated in the algorithm.  $G$  represents the neighborhood structure set included in the shaking process, where  $G = \{1, 2, 3, \dots, p, \dots, |G|\}$ , and  $H$  represents the neighborhood structure set included in the VND process, where  $H = \{1, 2, 3, \dots, p, \dots, |H|\}$ . The detailed structure of the improved VNS algorithm is shown in Algorithm 1.

The shaking process is a perturbation operator in the VNS. The process is used to generate different neighborhood solutions. The remaining initial solution and VND process are described in detail in the next section.

**4.1. Initial Solution.** The initial delivery route  $R$  of the MTS-OPH is formed by constructing a line of length  $2I'$  according to the number  $i'$  of a terminal cargo station. The initial route  $R$  is mainly composed of  $i'$  departure metro stations and  $i'$  arrival metro stations shown in Figure 4. Herein, each  $j_{i'}$  in  $[j_1, j_2, \dots, j_{i'-1}, j_{i'}]$  corresponds to each  $j'_{i'}$  in  $[j'_1, j'_2, \dots, j'_{i'-1}, j'_{i'}]$  individually, and a group  $[j_{i'}, j'_{i'}]$  constitutes a metro distribution route. The following conditions should be met:  $j \in J$ ,  $j' \in J'$ , and  $J = J'$ .

**4.2. Neighborhood Structure.** Two types of neighborhood structures are applied. The first consists of neighbors which exchange strategies on existing delivery routes. The second consists of neighbors obtained by updating the distribution strategy on the existing route. The methods for generating neighbors for the first type of neighborhood structure are

Swap-2 and Inserting-t [42]. For the second type of neighborhood structure, Alter-t method of producing neighbors is designed in our study.

Swap-2 refers to the random exchange of two adjacent or nonadjacent rows in the initial solution, as shown in Figure 5(a). Insertion-t is formed by repeating  $t$  times on the basis of Insertion-1. As shown in Figure 5(b), Insertion-1 randomly deletes a row from the initial solution and randomly inserts it into other positions. In Figure 5(c), Alter-1 randomly selects a position  $j_{i'}$  in the initial solution and selects a new number from the set  $J$  to replace  $j_{i'}$ . Alter-t is to repeat the Alter-1 operation  $t$  times. In addition, to prevent the value of  $t$  from being too large and destroying the stability of the obtained solution structure, the value of  $t$  is controlled to be in the range of  $[0, 5]$  in the insertion operation and change operation of this study. The iterations of repeated deletions, changes, and insertions let the algorithm search in a larger solution space, thereby enhancing the ability of the neighborhood search algorithm to move away from the local optimal region.

**4.3. Shaking Procedure.** We use  $G = \{\text{Alter-3}, \text{Alter-4}\}$  as the set of neighborhood structures in the shaking procedure. For each structure  $G_v \in G$ , it maps a given solution  $r$  to a series of neighborhoods  $G_v(c)$ . When the shaking procedure is applied, a solution will be randomly chosen from the neighborhoods. Accordingly, the detailed procedure is given in Algorithm 2.

**4.4. Variable Neighborhood Descent.** During a local search in VND, when a better solution than the current solution cannot be found in this neighborhood, the search is continued by moving to the next neighborhood solution. Contrarily, if a better solution than the current solution is found in this neighborhood, the first neighborhood solution will be returned to restart the search. For a better solution, the first neighborhood solution should be returned and the search should be started again. The local optimal solution obtained through such a search process is likely to be the global optimal solution. A detailed operation of the VND process is provided in Algorithm 3.

## 5. Numerical Experiments

Case studies of the Ningbo metro network and Beijing metro network were conducted to evaluate the accuracy and efficiency of the proposed model and method. Examples of the application in different delivery distances, different passenger flows, and different types of metro networks were illustrated. The proposed algorithm framework was coded in MATLAB 10.0 on a Window 10 personal computer with 4.0 GB processor. The MTS-OPH problem is solved by IBM CPLEX 12.5 Academic Version on the same platform.

**5.1. Small-Scale Case Study.** In this section, based on the logistics information of a certain express company in Ningbo, we considered a distribution situation between the

**Step 1:** a solution  $c$  is randomly generated in a given range, and the initial solution is brought into the objective function calculation program to obtain the initial target value  $f(c)$ ; go to Step 2.

**Step 2:** let  $p = 1$  until  $p = |G|$ ; go to Step 3.

**Step 3:** substitute the initial solution  $c$  into Shaking, and perform the  $k$ -th neighborhood operation of Shaking to generate a new solution  $c'$ ; go to Step 4.

**Step 4:** substitute  $c'$  into the VND, generate multiple neighborhood solutions related to  $c'$ , and perform a local search among the generated multiple neighborhood solutions to obtain the local search optimal solution  $c''$ ; go to Step 5.

**Step 5:** if  $f(c'') < f(c)$ , it means that  $c''$  is better than the current optimal solution  $c$ ; then, let  $c = c''$  and continue to search within the current neighborhood structure ( $p = 1$ ); otherwise, let  $p = p + 1$ . Meanwhile, if  $p = |G|$ , go to Step 6. Otherwise, go to Step 3.

**Step 6:** until the termination criterion is met, the optimal solution  $c$  is returned.

ALGORITHM 1: Procedure of VNS algorithm.

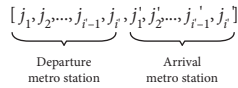


FIGURE 4: Initial path diagram.

express delivery distribution center and six terminal express delivery stations. For convenience, the cargo distribution center is represented by  $I$  and the six cargo stations are named A, B, C, D, E, and F, respectively, i.e.,  $i' = \{1, 2, 3, 4, 5, 6\}$ ,  $i' \in I'$ . The geographical location of freight stations is evenly distributed, and the receipt volume of each cargo station obeys the uniform distribution of [40, 60]. The metro network of the Ningbo Urban Rapid Rail Transit Construction Plan (2013–2020) [43] was considered as an example, as shown in Figure 6.

The airport logistics park near the metro line was selected as the express delivery distribution center, which is represented by a red star in Figure 6. L1–L5 indicate each metro line. Considering the actual situation and model solution, each metro line station was uniformly numbered according to positive integers from left to right, with the letter “I” representing the departure metro station and “O” representing the arrival metro station.

The average speed of freight vehicles was 20 km/h [30], and the delivery cost of unit express delivery per unit distance of freight vehicles was assumed to be 2 yuan [44]. The values of the other parameters are listed in Table 2. In the experiment, the distribution data and time window are shown in Table 3.

In the algorithm parameter setting process, this study sets the number of neighborhood solutions in one iteration to 100 and the total number of iterations to 50 [27]. The neighborhood structures of VND and shaking are as follows: VND, {Alter\_3, Alter\_5, Insertion\_2, Insertion\_1, Opt\_2}; Shaking, {Alter\_3, Alter\_5}.

For related problems in route planning and distribution, the genetic algorithm (GA) is widely used [45–47]. Therefore, the GA and the improved VNS algorithm are used to solve the problem. The population size of the GA is 100, the number of iterations is 50, the crossover probability is 0.6, and the mutation probability is 0.1 [48]. The optimization results of the two algorithms were obtained by running 20 times, as shown in Table 4. Although GA has a shorter calculating time, VNS can obtain the best delivery cost of higher quality, and

after many repeated trials, the results show that the stability of the optimal solution obtained by VNS is better than that of GA, which verifies the effectiveness of the algorithm.

In order to further prove the effectiveness of the improved VNS algorithm, we first give the actual maximum cross-section passenger flow to Equation (12), making MTS-OPH an integer linear programming model. Secondly, the ILOG CPLEX solver is used to solve it. After implementations, we finally obtained the returned solution with a calculation time of 328 s, where the relative gap turns out to be 5.00%. The approximate optimal objective value is 5861.0 yuan, which is consistent with the solution result of the improved VNS algorithm. Therefore, the proposed algorithm is effective.

Finally, the optimal solution for the delivery of vehicles is calculated. Meanwhile, the optimal transportation strategy during off-peak hours for the individual distribution of freight vehicles is calculated; this transportation strategy is called VTS-OPH. Table 5 shows the comparison between the MTS-OPH and VTS-OPH. The delivery time of VTS is mainly composed of two parts: loading and unloading time and transportation time. Among them, the loading and unloading time is determined according to the quantity of goods and the transportation time is determined by the transportation distance and the transportation speed. If the delivery time is less than 1 min, it is approximated to 1 min.

As evident from Table 5, the MTS-OPH and VTS-OPH complete deliveries within the time window. The average time to complete a delivery in the VTS-OPH is 69 min, and in the MTS-OPH, the average time is 73 min. However, the MTS-OPH spends approximately one-third of the delivery cost of the VTS-OPH to complete a delivery. According to this small-scale case analysis results, the MTS-OPH has certain advantages when compared with the VTS-OPH.

The delivery route is shown in Table 6, and the result shows that for the express station “F” delivery, the MTS-OPH requires less time than the VTS-OPH, and for the delivery of the other express stations, the VTS-OPH requires less time than the MTS-OPH. Therefore, further investigation is required regarding the applicability of the MTS-OPH for different distances.

*5.2. Applicability of MTS-OPH considering Different Delivery Distances.* To study the applicability of the MTS-OPH for different delivery distances, we divide the actual express

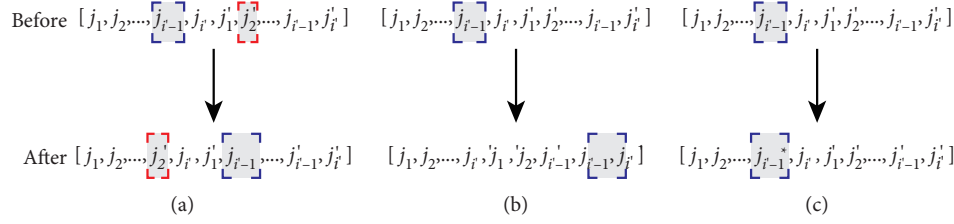


FIGURE 5: Three methods of producing neighbors. (a) Swap-2. (b) Insertion-1. (c) Alter-1.

**Step 1:** enter the initial solution  $c$  and the index of neighborhood structure  $G$ ; go to Step 2.  
**Step 2:** select a solution from  $G_v(c)$  at random, and then go to Step 3.  
**Step 3:** a random chosen solution  $c'$  is returned.

ALGORITHM 2: Shaking procedure.

**Step 1:** enter the initial solution  $c'$  and neighborhood structure set  $H$  of the VND; go to Step 2.  
**Step 2:** let  $p = 1$  until  $p = |H|$ , and then go to Step 3.  
**Step 3:** use the neighborhood structure  $H$  to perform a local search to generate a new solution  $c''$ , and then go to Step 4.  
**Step 4:** if  $f(c'') < f(c')$ , it means that  $c''$  is better than the current optimal solution  $c'$ ; then, let  $c' = c''$  and continue to search within the current neighborhood structure ( $p = 1$ ); otherwise, let  $p = p + 1$ . Meanwhile, if  $p = |H|$ , go to Step 5; otherwise, go to Step 3.  
**Step 5:** until the termination criterion is met, the optimal solution  $c'$  is returned.

ALGORITHM 3: Procedure of VND.

delivery information provided by the airport logistics park into three different distribution range data tables. The metro network of Ningbo is shown in Figure 7.

According to the metro network and express delivery information data, Ningbo is divided into three areas with different distribution scopes as S1, S2, and S3. Twenty delivery destinations (terminal express delivery stations) are selected with uniform locations in each area, as shown in Figure 8. Each delivery area is separated by a blue dotted line, where S1 is a short-distance delivery area, S2 is a medium-distance delivery area, and S3 is a long-distance delivery area. According to different distribution ranges, different time window requirements are allotted [44], as shown in Table 7.

In the experiment, the time horizon is considered to be 9:00–12:00, which is the off-peak period of metro operation. In addition, the metro model in Ningbo is type B composed of six carriages with a total capacity of 1,460 people. According to statistics from the Ningbo Rail Transit Group, the maximum load rate during off-peak hours in the Ningbo metro network is 90%. Hence, at least five metro carriages are required for passenger transportation during off-peak hours, and the remaining one carriage is used for goods transportation. The remaining parameter values and algorithm parameter settings are consistent with the small-scale case study presented in Section 5.1.

**5.2.1. Analysis of the Short-Distance Delivery Area.** The demand and time window of the terminal express station in S1 are listed in Table 8.

The data in Table 8 are used in the VNS algorithm for 20 calculations to obtain the optimal solution for the MTS-OPH and VTS-OPH, as shown in Table 9, and the optimal delivery routes of the two strategies are given in Table 10.

In Table 9, the delivery cost of the MTS-OPH is approximately 66.4% higher than that of the VTS-OPH. In addition, 12 of the 20 terminal express delivery stations in the MTS-OPH failed to deliver on time, while the VTS-OPH completed all deliveries with the set time window. In the short-distance distribution area, the success rate of the MTS-OPH to complete the delivery within the time window was 40%. However, the MTS-OPH reduces the vehicle delivery distance by 56.7% when compared with that of the VTS-OPH. Although the vehicle transportation distance is reduced, the delivery time is increased, as shown in Table 10. According to the above analysis, in the short-distance delivery area, the MTS-OPH is not suitable for multitarget delivery and the overall operation of express delivery companies, but the single-target delivery remains to be studied. This corresponds well with the study of high-speed railway freight distribution by Pazour et al. [49].

**5.2.2. Analysis of the Medium-Distance Delivery Area.** The demand and time window of the terminal express station in S2 are listed in Table 11.

Similarly, the data in Table 11 are used in the VNS algorithm for 20 calculations, and the optimal solution is listed in Table 12.

As shown in Table 12, the delivery cost of the VTS-OPH is 8.0% higher than that of the MTS-OPH, but the average delivery time of the VTS-OPH is 33.7% less than that of the



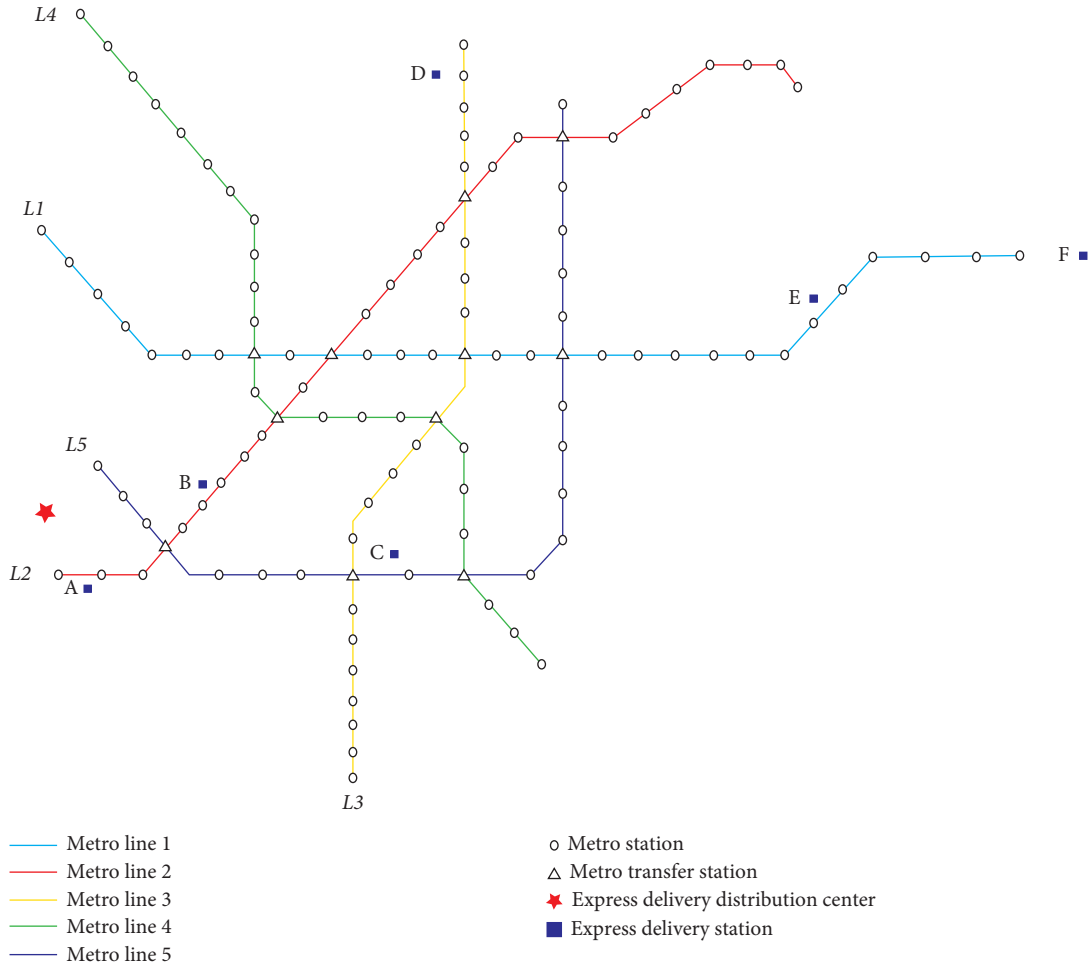


FIGURE 6: Ningbo metro network schematic diagram.

TABLE 2: Values and units of certain parameters in the small-scale case study.

Parameters	Notations	Values
Vehicle fixed operating cost	$C_v$	30 (yuan)
Capacity of the transport vehicle	$cap_V$	60 (parcel)
Freight capacity of one metro carriage	$cap_M$	330 (parcel)
Number of freight carriages	$E$	1 (carriage)
Transfer cost	$C_O$	0.6 yuan/parcel
Transfer time	$t_O$	6 min
Headway	$t_{interval}$	7 min
The delay cost	$\alpha$	2 yuan/parcel

TABLE 3: Express delivery station demand and time window in the small-scale case study.

Express delivery station	Express delivery station demand (parcel)	Time window (min)
A	53	[0, 60]
B	55	[0, 180]
C	59	[0, 60]
D	57	[0, 180]
E	51	[0, 180]
F	58	[0, 180]



TABLE 4: Comparison of optimization results for two algorithms.

Algorithm	Optimal calculating time (s)	Best delivery cost (yuan)	Worst delivery cost (yuan)	Average delivery cost (yuan)
VNS	86.3	5861.0	6239.8	6021.2
GA	16.2	6382.5	7618.2	6775.0

TABLE 5: Optimal solution for the two transportation strategies.

Transportation strategy	Delivery cost (yuan)	Delivery distance by vehicles (km)	Average delivery time (min)	Number of express delivery stations that did not deliver on time
MTS-OPH (Method 1)	5861.0	26.3	73	0
VTS-OPH	15350.0	136.4	69	0

TABLE 6: Delivery route comparison.

Terminal express stations	MTS-OPH (Method 1)		VTS-OPH	
	Delivery route (transfer times)	Delivery time (min)/train number	Delivery route (transfer times)	Delivery time (min)
A	I → L5I2 → L2O2 → A (1)	39/1	I → A (0)	21
B	I → L5I2 → L2O6 → B (1)	41/1	I → B (0)	20
C	I → L5I2 → L5O9 → C (0)	41/1	I → C (0)	36
D	I → L5I2 → L3O22 → C (1)	87/1	I → D (0)	75
E	I → L5I2 → L1O24 → C (1)	101/1	I → E (0)	111
F	I → L5I2 → L1O29 → C (1)	125/2	I → F (0)	147

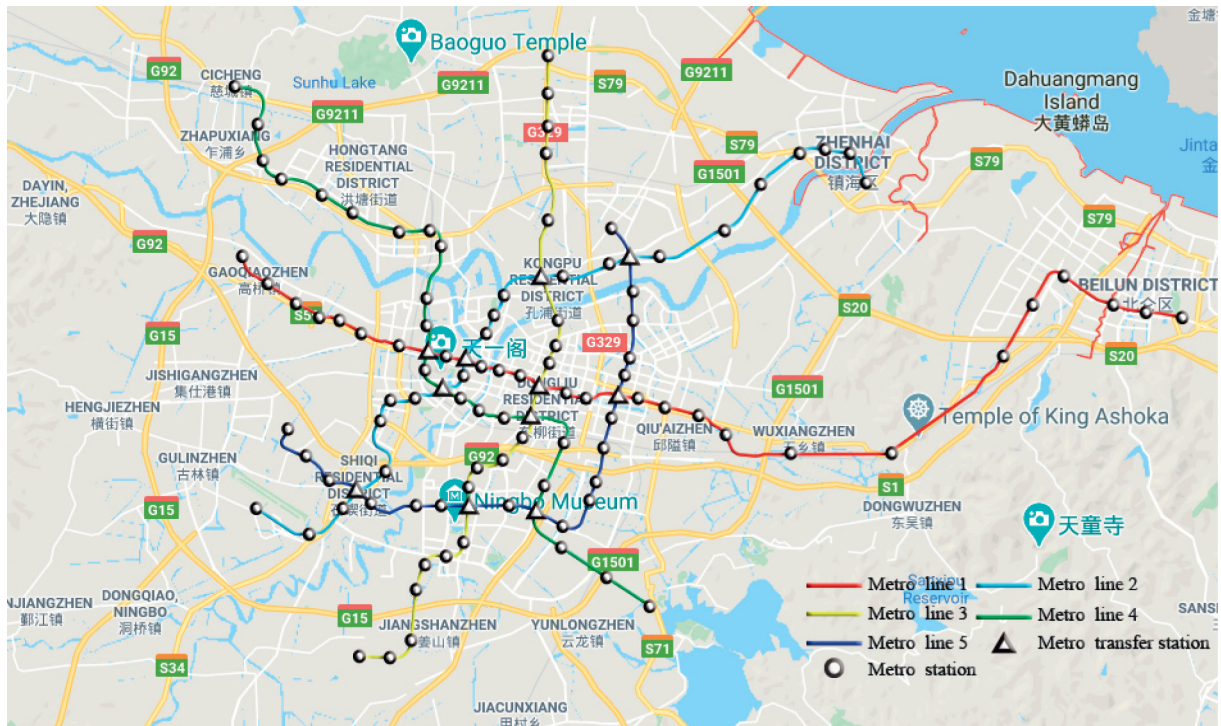


FIGURE 7: Metro network of Ningbo in 2020.

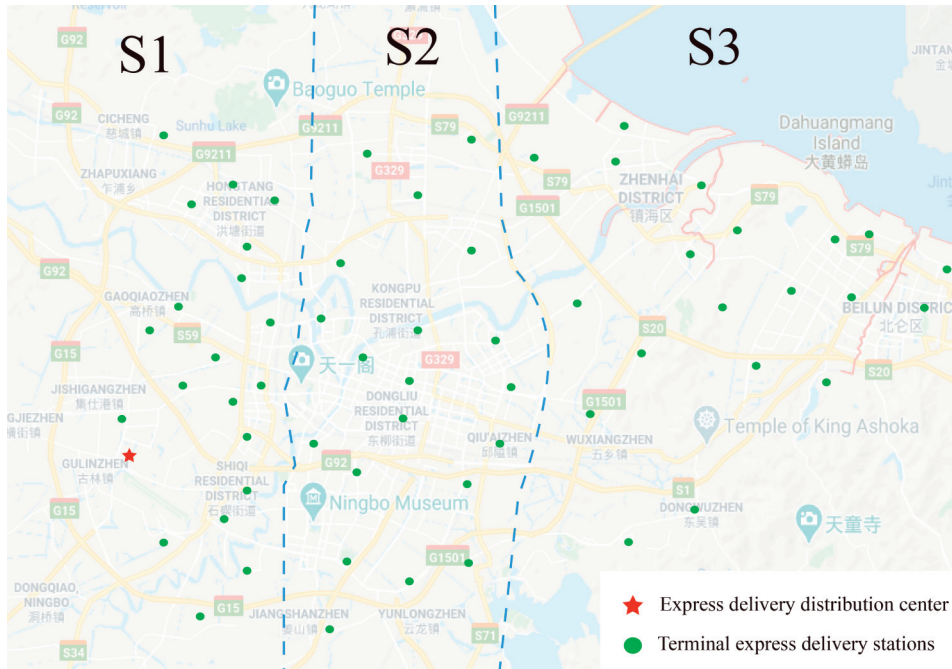


FIGURE 8: Distribution area.

TABLE 7: Delivery time window.

Delivery area	S1	S2	S3
Delivery time window [ET, LT] (min)	[0, 60] or [0, 90]	[0, 90] or [0, 120]	[0, 120] or [0, 180]

TABLE 8: Express delivery station demand and time window in S1.

Express delivery station	Express delivery station demand (parcel)	Time window (min)
1	56	[0, 90]
2	58	[0, 60]
3	43	[0, 90]
4	58	[0, 90]
5	53	[0, 90]
6	42	[0, 90]
7	46	[0, 90]
8	51	[0, 60]
9	59	[0, 90]
10	59	[0, 60]
11	43	[0, 90]
12	59	[0, 60]
13	59	[0, 60]
14	50	[0, 60]
15	56	[0, 60]
16	43	[0, 90]
17	48	[0, 90]
18	58	[0, 60]
19	56	[0, 90]
20	59	[0, 60]

TABLE 9: Optimal solution for two transportation strategies.

Transportation strategy	Delivery cost (yuan)	Delivery distance by vehicles (km)	Average delivery time (min)	Number of express delivery stations that did not deliver on time
MTS-OPH (Method 1)	38842.3	93.3	83	12
VTS-OPH	23349.0	214.8	33	0

TABLE 10: Delivery route comparison.

Terminal express stations	MTS-OPH (Method 1)		VTS-OPH	
	Delivery route (transfer times)	Delivery time (min)/train number	Delivery route (transfer times)	Delivery time (min)
1	I → L5I2 → L4O2 → 1 (2)	109/2	I → 1 (0)	72
2	I → L5I2 → L4O3 → 2 (2)	97/1	I → 2 (0)	57
3	I → L5I2 → L4O5 → 3 (2)	108/2	I → 3 (0)	54
4	I → L5I2 → L4O6 → 4 (2)	87/3	I → 4 (0)	57
5	I → L2I1 → L4O6 → 5 (1)	90/1	I → 5 (0)	45
6	I → L5I2 → L4O6 → 6 (2)	102/3	I → 6 (0)	42
7	I → L5I2 → L1O3 → 7 (2)	97/3	I → 7 (0)	28
8	I → L5I2 → L1O3 → 8 (2)	88/1	I → 8 (0)	23
9	I → L5I2 → L1O7 → 9 (2)	92/3	I → 9 (0)	33
10	I → L5I2 → L1O7 → 10 (2)	84/1	I → 10 (0)	22
11	I → L5I2 → L1O6 → 11 (2)	87/3	I → 11 (0)	27
12	I → L5I2 → L2O8 → 12 (1)	49/1	I → 12 (0)	19
13	I → L5I2 → L1O5 → 13 (2)	85/1	I → 13 (0)	15
14	I → L5I2 → L1O2 → 14 (2)	93/2	I → 14 (0)	11
15	I → L5I2 → L2O7 → 15 (1)	48/2	I → 15 (0)	21
16	I → L5I2 → L2O4 → 16 (1)	42/3	I → 16 (0)	18
17	I → L5I2 → L2O3 → 17 (1)	75/4	I → 17 (0)	17
18	I → L5I2 → L2O2 → 18 (1)	50/2	I → 18 (0)	20
19	I → L5I2 → L3O1 → 19 (1)	93/4	I → 19 (0)	39
20	I → L5I2 → L3O5 → 20 (1)	67/2	I → 20 (0)	29

TABLE 11: Express delivery station demand and time window in S2.

Express delivery station	Express delivery station demand (parcel)	Time window (min)
1	47	[0, 90]
2	57	[0, 90]
3	52	[0, 120]
4	51	[0, 90]
5	58	[0, 120]
6	46	[0, 90]
7	55	[0, 90]
8	55	[0, 120]
9	48	[0, 120]
10	51	[0, 120]
11	42	[0, 120]
12	41	[0, 120]
13	51	[0, 120]
14	56	[0, 120]
15	59	[0, 90]
16	43	[0, 120]
17	51	[0, 90]
18	49	[0, 90]
19	40	[0, 90]
20	47	[0, 90]

TABLE 12: Optimal solution for two transportation strategies.

Transportation strategy	Delivery cost /(yuan)	Delivery distance by vehicles (km)	Average delivery time (min)	Number of express delivery stations that did not deliver on time
MTS-OPH (Method 1)	35582.5	94.1	86	2
VTS-OPH	38661.6	378.4	57	0

MTS-OPH. In addition, the specific delivery time, delivery route, transfer times, and train number of each express station under the two modes are shown in Table 13. Concurrently, VTS-OPH completed all deliveries within the time window. There were two express delivery stations in the MTS-OPH that failed to deliver on time. The vehicle delivery distance in MTS-OPH is 24.7% of that in VTS-OPH.

Combining the results listed in Tables 9 and 12, the number of express stations that did not deliver on time significantly improved when the MTS-OPH was selected in the medium-distance delivery area, when compared with the MTS-OPH in the short-distance delivery area. From the original 12 express stations that failed to deliver on time, two express delivery stations failed to deliver on time. Therefore, although the VTS-OPH requires less time to complete 100% on-time distribution, the MTS-OPH delivery cost is slightly lower and the MTS-OPH significantly reduces the vehicle transportation distance. This means that when choosing the MTS-OPH for delivery, the transportation distance of vehicles can be reduced and pressure on urban roads can be reduced. According to the above analysis, the MTS-OPH is more suitable for the medium-distance distribution area than for the short-distance distribution area. The MTS-OPH and VTS-OPH exhibit varied advantages in medium-distance delivery areas; however, considering the urban system, the MTS-OPH is better than the VTS-OPH.

*5.2.3. Analysis of the Long-Distance Delivery Area.* The demand and time window of the terminal express station in S3 are shown in Table 14.

According to the above experimental data, the VNS algorithm is used to solve the two transportation strategies of MTS-OPH and VTS-OPH 20 times and the optimal solution obtained is shown in Table 15. The delivery route information is shown in Table 16.

As shown in Tables 15 and 16, we obtained a different delivery result when compared with results of previous analyses. In the long-distance delivery area, all indicators of the MTS-OPH are better than those of the VTS-OPH, except for the average delivery time. The delivery cost of MTS-OPH is 52.2% of VTS-OPH, the vehicle delivery distance is 17.6% of VTS-OPH, and the average delivery time is slightly worse than that of VTS-OPH. A significant change is observed in the number of express delivery stations that failed to deliver on time. In the MTS-OPH, all express delivery stations completed the delivery within the time window, while in the VED mode, there were four express stations that failed to complete the delivery within the time window. Thus, the MTS-OPH is suitable for long-distance distribution areas. A similar result was reported in a high-speed rail express delivery study conducted in [50].

Concurrently, by comparing the calculation results in Tables 9, 12, and 15, it is further verified that the MTS-OPH is suitable for long-distance delivery. On the one hand, as the distribution distance increased, the distribution cost of the MTS-OPH changed slightly and the vehicle delivery distance increased at an average growth rate of 18%, which was much smaller than that of the VTS-OPH and increased with an

average growth rate of 86%. The above analysis proves the feasibility and stability of the MTS-OPH. On the other hand, from short-distance delivery to medium-distance delivery to long-distance delivery, the number of express stations that failed to complete the delivery within the time window under the MTS-OPH changed from twelve to zero and that of the VTS-OPH changed from zero to four.

*5.2.4. Relationship between MTS-OPH and VTS-OPH under Different Delivery Distances.* To further explore the relationship between the MTS-OPH and VTS-OPH under different delivery distances, we assumed that there was one terminal express station, did not consider the time window constraints, and quantitatively analyzed the delivery costs of the two transportation strategies.

$$f_c^m = Q * (2 * (L_{ij} + L_{j'i'}) + C_o * (O_{jj'} + 2)) + 2 * \frac{Q}{cap_v} * C_v, \quad (28)$$

$$f_c^v = 2Q * L_{ii'} + \frac{Q}{cap_v} * C_v, \quad (29)$$

where  $f_c^m$  represents the delivery cost of MTS-OPH,  $f_c^v$  represents the delivery cost of VTS-OPH,  $Q$  is the number of express parcels,  $C_v$  is the fixed operating cost of the vehicle,  $cap_v$  is the capacity of freight vehicles,  $L$  is the vehicle transportation distance between two points,  $C_o$  is per unit cargo transfer cost after one transfer, and  $O$  is the number of transfers between two stations.

The delivery costs of the two transportation strategies are shown in Equations (28) and (29). Assuming that the number of express deliveries for delivery is 60 parcels that reach the upper limit of vehicle loading and the number of transfers is 2, the relationship between the two transportation strategies is

$$L_{ii'} > 1.45 + (L_{ij} + L_{j'i'}). \quad (30)$$

Equation (30) shows that when the vehicle delivery distance of VTS-OPH is greater than the sum of 1.45 and the vehicle delivery distance of MTS-OPH, the MTS-OPH should be adopted; otherwise, the VTS-OPH should be selected. The above conclusions further prove that the MTS-OPH is more suitable for long-distance delivery.

*5.3. Impact on MTS-OPH under Different Load Rates of Passenger Flow during Off-Peak Hours.* In this section, we analyze the impact of different load rates of passenger flow on MTS-OPH. We selected case data in S3 for analysis, the total number of express deliveries is 975 packages, and three train carriages are required as freight carriages. This determines the 50% full-load rate during off-peak hours as the boundary. Concurrently, following the principle of passenger and cargo diversion, according to the statistics of Ningbo Rail Transit Group, the maximum passenger flow load rate of the Ningbo metro network during off-peak hours is 90% and 5 train carriages are required as passenger

TABLE 13: Delivery route comparison.

Terminal express stations	MTS-OPH (Method 1)		VTS-OPH	
	Delivery route (transfer times)	Delivery time (min)/train number	Delivery route (transfer times)	Delivery time (min)
1	I → L5I2 → L3O22 → 1 (2)	94/1	I → 1 (0)	72
2	I → L5I2 → L3O23 → 2 (2)	105/1	I → 2 (0)	86
3	I → L5I2 → L3O21 → 3 (2)	101/2	I → 3 (0)	72
4	I → L5I2 → L5O22 → 4 (1)	87/1	I → 4 (0)	69
5	I → L5I2 → L3O18 → 5 (2)	94/2	I → 5 (0)	60
6	I → L5I2 → L3O18 → 6 (2)	85/1	I → 6 (0)	45
7	I → L5I2 → L2O15 → 7 (1)	70/1	I → 7 (0)	45
8	I → L5I2 → L1O18 → 8 (1)	97/3	I → 8 (0)	69
9	I → L5I2 → L3O18 → 9 (2)	104/3	I → 9 (0)	51
10	I → L5I2 → L5O20 → 10 (1)	97/3	I → 10 (0)	72
11	I → L5I3 → L4O19 → 11 (1)	76/3	I → 11 (0)	51
12	I → L5I2 → L1O20 → 12 (1)	102/3	I → 12 (0)	60
13	I → L5I2 → L2O16 → 13 (1)	91/3	I → 13 (0)	36
14	I → L5I2 → L4O17 → 14 (1)	85/4	I → 14 (0)	39
15	I → L5I2 → L3O11 → 15 (1)	54/1	I → 15 (0)	60
16	I → L5I2 → L1O19 → 16 (1)	106/4	I → 16 (0)	39
17	I → L5I2 → L3O3 → 17 (1)	75/2	I → 17 (0)	51
18	I → L5I2 → L3O6 → 18 (1)	62/2	I → 18 (0)	57
19	I → L5I2 → L4O24 → 19 (1)	70/2	I → 19 (0)	54
20	I → L5I2 → L4O23 → 20 (1)	62/2	I → 20 (0)	48

TABLE 14: Express delivery station demand and time window in S3.

Express delivery station	Express delivery station demand (parcel)	Time window (min)
1	56	[0, 120]
2	48	[0, 120]
3	45	[0, 180]
4	48	[0, 180]
5	42	[0, 120]
6	43	[0, 120]
7	59	[0, 180]
8	59	[0, 120]
9	52	[0, 120]
10	41	[0, 120]
11	45	[0, 120]
12	47	[0, 180]
13	56	[0, 120]
14	40	[0, 120]
15	41	[0, 180]
16	43	[0, 120]
17	53	[0, 120]
18	55	[0, 180]
19	53	[0, 180]
20	49	[0, 180]

TABLE 15: Optimal solution for two transportation strategies.

Transportation strategy	Delivery cost (yuan)	Delivery distance by vehicles (km)	Average delivery time (min)	Number of express delivery stations that did not deliver on time
MTS-OPH (Method 1)	38131.8	130.8	115	0
VTS-OPH	73006.8	743.6	112	4

carriages. Therefore, 90% of the passenger flow load rate is determined as upper limit. When the full-load rate of passenger flow during the off-peak hours is greater than

90%, the MTS-OPH is not selected, as shown in Figure 9. When the interval for the full-load rate of passenger flow is [0%, 50%] and the total number of goods and passengers is



TABLE 16: Delivery route comparison.

Terminal express stations	MTS-OPH (Method 1)		VTS-OPH	
	Delivery route (transfer times)	Delivery time (min)/train number	Delivery route (transfer times)	Delivery time (min)
1	I → L5I2 → L2O24 → 1 (1)	105/1	I → 1 (0)	93
2	I → L5I2 → L2O26 → 2 (1)	112/1	I → 2 (0)	114
3	I → L5I2 → L2O26 → 3 (1)	127/2	I → 3 (0)	102
4	I → L5I2 → L2O24 → 4 (1)	129/3	I → 4 (0)	108
5	I → L2I1 → L2O26 → 5 (1)	112/1	I → 5 (0)	108
6	I → L5I2 → L2O21 → 6 (1)	103/1	I → 6 (0)	87
7	I → L5I2 → L1O22 → 7 (1)	118/3	I → 7 (0)	87
8	I → L5I2 → L1O21 → 8 (1)	91/1	I → 8 (0)	78
9	I → L5I2 → L1O23 → 9 (1)	96/1	I → 9 (0)	93
10	I → L5I2 → L1O27 → 10 (1)	118/2	I → 10 (0)	126
11	I → L5I2 → L1O24 → 11 (1)	101/2	I → 11 (0)	108
12	I → L5I2 → L1O25 → 12 (1)	115/3	I → 12 (0)	120
13	I → L5I2 → L2O27 → 13 (1)	119/2	I → 13 (0)	114
14	I → L5I2 → L1O26 → 14 (1)	114/2	I → 14 (0)	126
15	I → L5I2 → L1O26 → 15 (1)	127/3	I → 15 (0)	135
16	I → L5I2 → L1O26 → 16 (1)	108/2	I → 16 (0)	126
17	I → L5I2 → L1O28 → 17 (1)	116/2	I → 17 (0)	147
18	I → L5I2 → L1O28 → 18 (1)	123/3	I → 18 (0)	147
19	I → L5I2 → L1O26 → 19 (1)	128/3	I → 19 (0)	120
20	I → L5I2 → L1O22 → 20 (1)	119/4	I → 20 (0)	92

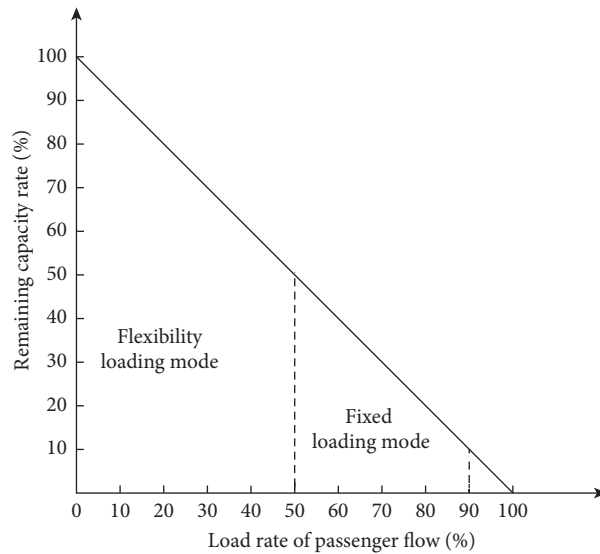


FIGURE 9: Transportation methods under different load rates of passenger flow.

less than the capacity of the metro train, the flexibility loading mode should be selected, the full-load rate of passenger flow should be (50%, 90%], the total number of goods and passengers is greater than the capacity of the metro train, and the fixed loading mode should be selected. When the full-load rate of passenger flow is greater than 90%, the metro is used for transporting passengers.

To further compare the two transportation methods in the MTS-OPH, with the load rates of passenger flow of 50% and 90% as the boundary, the VNS algorithm was run 20 times to obtain the optimal results, as shown in Table 17.

In Table 17, the delivery cost and the vehicle delivery distance are equal in Methods 1 and 2, but considering the average delivery time, Method 2 is shorter than Method 1. When the load rate of passenger flow satisfies the boundary conditions, the delivery efficiency of Method 2 is better than that of Method 1 for the same delivery time. Considering long-term transportation, MTS-OPH has the potential and has huge positive effects; for example, it can alleviate traffic congestion, has huge economies of scale [51, 52], and has much lower fuel consumption than the VTS-OPH. The MTS-OPH is adopted during off-peak hours, which can



TABLE 17: Optimal solution for the two transportation methods of MTS-OPH.

Transportation method of MTS-OPH	Delivery cost (yuan)	Delivery distance by vehicles (km)	Average delivery time (min)	On-time delivery rate (%)
Method 1	37852.2	130.8	114	100
Method 2	37852.2	130.8	106	100

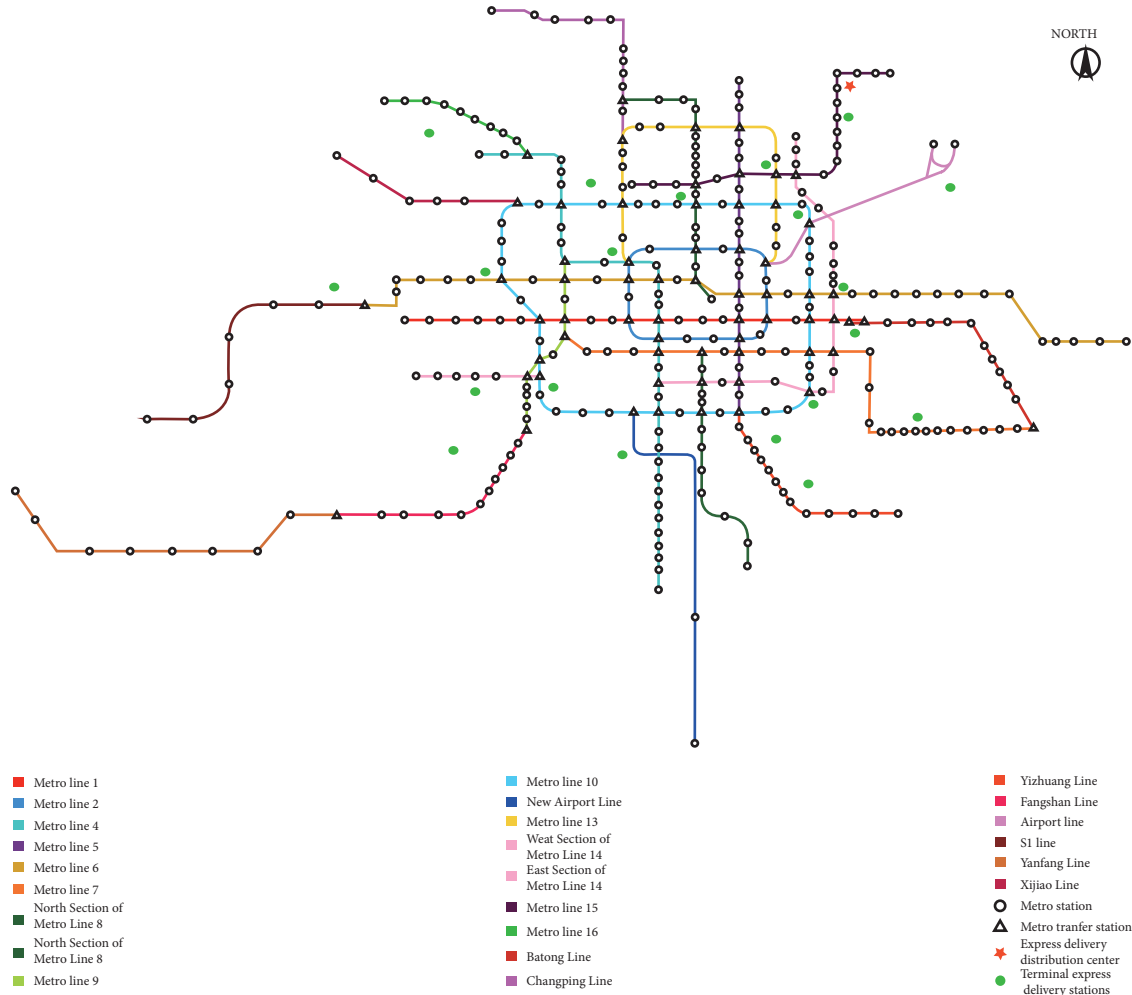


FIGURE 10: Metro network of Beijing.

fulfill the use of the idle resources of the metro and does not require excessive initial investment.

**5.4. Applicability of MTS-OPH considering Different Metro Networks.** In order to further prove the applicability and feasibility of the model, we selected the Beijing metro network for numerical experiments. The Beijing metro network is one of the most complex metro networks in China, with 24 metro lines totaling 331 stations, including 62 transfer stations, as shown in Figure 10. In Figure 10, the green dots indicate the terminal express delivery stations, totaling 20. The red five-pointed star indicates the express delivery distribution center, which is the airport logistics park near

the capital airport. Different from Ningbo’s radial metro network, Beijing’s metro network is a ring-shaped radial network with more diversified route selections between lines. Therefore, we increased the maximum number of transfers in the model to 4 times.

In the experiment, the considered time horizon is set as 9:00–12:00, which is the off-peak period of metro operation. According to the data of Beijing Subway Operation Company, the maximum passenger load rates of passenger flow in the Beijing metro network during off-peak hours is 94.3%. The metro train is composed of 6 carriages, and it needs to occupy 5 carriages for passenger transportation. Thus, the loading mode of Method 1 for mixed transportation is selected. Some parameters in the experiments

TABLE 18: Express delivery station demand and time window in Beijing.

Express delivery station	Express delivery station demand (parcel)	Time window (min)
1	56	[0, 120]
2	58	[0, 120]
3	43	[0, 180]
4	58	[0, 180]
5	53	[0, 120]
6	42	[0, 180]
7	46	[0, 180]
8	51	[0, 120]
9	59	[0, 120]
10	59	[0, 180]
11	43	[0, 180]
12	59	[0, 180]
13	59	[0, 120]
14	50	[0, 180]
15	56	[0, 180]
16	43	[0, 120]
17	48	[0, 180]
18	58	[0, 120]
19	56	[0, 180]
20	59	[0, 120]

TABLE 19: Optimal solution for two transportation strategies.

Transportation strategy	Delivery cost (yuan)	Delivery distance by vehicles (km)	Average delivery time (min)	Number of express delivery stations that did not deliver on time
MTS-OPH (Method 1)	49320.8	265.0	104	0
VTS-OPH	82855.2	776.0	117	3

TABLE 20: Delivery route comparison.

Terminal express stations	MTS-OPH (Method 1)		VTS-OPH	
	Delivery route (transfer times)	Delivery time (min)/ train number	Delivery route (transfer times)	Delivery time (min)
1	I → L15I3 → L15O7 → 1 (0)	46/1	I → 1 (0)	27
2	I → L15I3 → L15O11 → 2 (0)	72/1	I → 2 (0)	75
3	I → LAI1 → L15O7 → 3 (1)	81/1	I → 3 (0)	45
4	I → L15I3 → L10O12 → 4 (1)	77/2	I → 4 (0)	84
5	I → L15I3 → L14O10 → 5 (1)	82/1	I → 5 (0)	91
6	I → L15I3 → L16O8 → 6 (4)	163/2	I → 6 (0)	138
7	I → L15I3 → L13O14 → 7 (1)	83/2	I → 7 (0)	112
8	I → L15I3 → L15O18 → 8 (0)	78/1	I → 8 (0)	99
9	I → L15I3 → L4O12 → 9 (2)	108/1	I → 9 (0)	106
10	I → LAI1 → L6O26 → 10 (1)	117/1	I → 10 (0)	129
11	I → L14EI1 → LS1O3 → 11 (2)	173/2	I → 11 (0)	182
12	I → L15I3 → L14EO5 → 12 (3)	164/1	I → 12 (0)	155
13	I → LAI1 → L9O11 → 13 (1)	98/3	I → 13 (0)	172
14	I → L15I3 → L1O22 → 14 (2)	28/3	I → 14 (0)	147
15	I → L15I3 → L4O24 → 15 (2)	149/1	I → 15 (0)	154
16	I → L14EI1 → L14EO16 → 16 (0)	115/1	I → 16 (0)	115
17	I → L5I3 → L2O25 → 17 (1)	155/2	I → 17 (0)	117
18	I → L15I3 → LYZO8 → 18 (2)	39/3	I → 18 (0)	143
19	I → L15I3 → L7O9 → 19 (2)	137/2	I → 19 (0)	143
20	I → L15I3 → L1O2 → 20 (2)	110/2	I → 20 (0)	102

are set as follows. That is, the delivery time window is set to 2 hours or 3 hours, and the demand and time windows of each terminal express station are shown in Table 18. The

remaining parameter values and algorithm parameter settings are consistent with the small-scale case study presented in Section 5.1.

The experimental data in Table 18 are used in the VNS algorithm for 20 calculations. The optimal solution and delivery route information are listed in Tables 19 and 20, respectively.

In Table 19, the delivery cost of MTS-OPH is 59.5% of VTS-OPH and the vehicle transportation distance is 34.1% of VTS-OPH. The average delivery time of MTS-OPH is slightly better than that of VTS-OPH, and the delivery tasks are all completed within the time window. However, under the VTS-OPH mode, there are 3 terminal express stations that failed to complete the delivery within the time window. In Table 20, terminal express stations closer to the express distribution center use VTS-OPH for shorter delivery time, while terminal express stations farther away from the express distribution center use MTS-OPH for shorter delivery time. In general, the MTS-OPH mode can complete the delivery within the time window, and the average delivery time is shorter, which is better than the VTS-OPH mode.

According to the case analysis results in Tables 19 and 20, MTS-OPH is more suitable for long-distance multitarget delivery. Compared with VTS-OPH, MTS-OPH has lower total delivery cost, shorter vehicle transportation distance, and higher service level. The above analysis results are consistent with the Ningbo metro case analysis results, which prove that our proposed model can be applied to different types of metro networks with different levels of complexity.

## 6. Conclusion

This paper proposes a new mixed transport strategy based on the metro network during off-peak hours to determine the freight mode and its distribution cost under off-peak metro passenger flow. The mixed transportation standard of passenger flow priority and separate transportation of passenger and cargo flows of the same train are proposed. According to the aforementioned criteria, a nonlinear programming model of the mixed transport strategy is constructed. In addition, an improved VNS algorithm is designed to solve the model. Finally, considering the Ningbo and Beijing metro network as an example, it is verified that the proposed model and mixed transport strategy can provide decision support for logistics companies. The main contributions of the study are as follows.

A theoretical framework of the mixed transport strategy for different metro passenger flows during off-peak hours was developed via comparisons with related studies [9, 20, 29]. Thus, a mixed transportation standard with passenger flow priority and separate transportation of the same train passenger and cargo flows was established, which expanded the integrated transportation of urban freight and metro passenger transportation.

In practice, the model of the mixed transportation strategy proposed in this study can provide decision support for logistics companies based on different delivery distances, different off-peak passenger flows, and different types of metro networks. First, when the vehicle transportation distance under separate VTS-OPH is greater than the sum of 1.45 and the vehicle transportation distance of MTS-OPH,

the MTS-OPH should be adopted; otherwise, the VTS-OPH should be selected. Second, when the number of passengers and freight transportation demand are greater than the metro capacity, the fixed loading mode of the MTS-OPH should be selected for transportation; otherwise, the flexibility loading mode of the MTS-OPH should be selected for transportation. Finally, the proposed model framework can be applied to different types of metro networks with different levels of complexity.

This study has some limitations to be further solved, such as freight vehicle scheduling and route planning and metro station and line performance evaluation [53, 54]. The study should mainly focus on the following aspects: (1) Freight vehicle routes should be planned to improve vehicle utilization, and train schedule issues should be analyzed from a data-driven perspective [55]. (2) Owing to the uneven distribution of passenger flow in the urban metro system during off-peak hours, the remaining loading capacity of other carriages can be considered in future studies. (3) To ensure the mixed strategy realizable, the metro turnover and rolling stock circulation and selecting the unloading stations should be considered.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Acknowledgments

This work was financially supported by the National Natural Science Foundation of China under Grants 71662011 and 71940009. This research was also jointly supported by Jiangxi Provincial Major Science and Technology Project-5G Research Project (Grant no. 20212ABC03A07).

## References

- [1] N. Keya, S. Anowar, and N. Eluru, "Joint model of freight mode choice and shipment size: a copula-based random regret minimization framework," *Transportation Research Part E: Logistics and Transportation Review*, vol. 125, pp. 97–115, 2019.
- [2] M. Lafkihi, S. Pan, and E. Ballot, "Freight transportation service procurement: a literature review and future research opportunities in omnichannel E-commerce," *Transportation Research Part E: Logistics and Transportation Review*, vol. 125, pp. 348–365, 2019.
- [3] M. Lindholm and S. Behrends, "Challenges in urban freight transport planning—a review in the Baltic Sea Region," *Journal of Transport Geography*, vol. 22, pp. 129–136, 2012.
- [4] National Institute of Statistics and Economic Studies, *The National Institute of Statistics and Economics Collects, Produces, Analyzes and Disseminates Information on the Economy and the French Economy*, National Institute of Statistics and Economic Studies, Paris, France, 2015.

- [5] J. Tang, A. McNabola, B. Misstear, and B. Caulfield, "An evaluation of the impact of the Dublin port tunnel and HGV management strategy on air pollution emissions," *Transportation Research Part D: Transport and Environment*, vol. 52, pp. 1–14, 2017.
- [6] The Central People's Government of the People's Republic of China, *Outline of Building a Powerful Transportation Country*, The Central People's Government of the People's Republic of China, Beijing, China, 2018.
- [7] China Urban Rail Transit Association, *Urban Rail Transit 2019 Annual Statistics and Analysis Report*, China Urban Rail Transit Association, Beijing, China, 2018.
- [8] W. Behiri, S. Belmokhtar-Berraf, and C. Chu, "Urban freight transport using passenger rail network: scientific issues and quantitative analysis," *Transportation Research Part E: Logistics and Transportation Review*, vol. 115, pp. 227–245, 2018.
- [9] J. Kikuta, T. Ito, I. Tomiyama, S. Yamamoto, and T. Yamada, "New subway-integrated city logistics system," *Procedia-Social and Behavioral Sciences*, vol. 39, pp. 476–489, 2012.
- [10] D. Reece and M. Marinov, "Modelling the implementation of a baggage transport system in Newcastle upon Tyne for passengers using mixedmode travel," *Transport Problem*, vol. 10, no. 4, pp. 149–155, 2015.
- [11] C. A. Nash, *Economics of Public Transport*, Longman, London, UK, 1982.
- [12] A. Trentini, R. Masson, F. Lehuédé, N. Malhéné, O. Péton, and H. Tlahig, *A Shared "Passengers & Goods" City Logistics System*, ILS, Quebec, Canada, 2012.
- [13] A. Trentini and N. Malhene, "Flow management of passengers and goods coexisting in the urban environment: conceptual and operational points of view," *Procedia-Social and Behavioral Sciences*, vol. 39, pp. 807–817, 2012.
- [14] J. Gonzalez-Feliu, *Costs and Benefits of Railway Urban Logistics: A Prospective Social Cost Benefit Analysis*, HAL, France, 2014.
- [15] E. Fatnassi, J. Chaouachi, and W. Klibi, "Planning and operating a shared goods and passengers on-demand rapid transit system for sustainable city-logistics," *Transportation Research Part B: Methodological*, vol. 81, pp. 440–460, 2015.
- [16] R. Masson, A. Trentini, F. Lehuédé, N. Malhéné, O. Péton, and H. Tlahig, "Optimization of a city logistics transportation system with mixed passengers and goods," *EURO Journal on Transportation and Logistics*, vol. 6, no. 1, pp. 81–109, 2017.
- [17] H. Quak, *Sustainability of Urban Freight Transport: Retail Distribution and Local Regulations in Cities*, Erasmus University Rotterdam, Rotterdam, Netherlands, 2008.
- [18] L. Dablaner, *Freight Transport for Development Toolkit: Urban Freight*, The World Bank, Washington, DC, USA, 2009.
- [19] Metropolitan Transportation Authority, *NYCT Trash Can Free Stations Pilot Update*, Metropolitan Transportation Authority, New York, NY, USA, 2020.
- [20] J. C. Rijsenbrij and B. A. Pielage, *The Potential of Metro Systems for City Logistics, Underground Freight Transportation by Capsule Pipelines and Other Tube/Tunnel Systems*, Ruhr-Universität Bochum, Bochum, Germany, 2002.
- [21] A. Motraghi and M. V. Marinov, "Analysis of urban freight by rail using event based simulation," *Simulation Modelling Practice and Theory*, vol. 25, pp. 73–89, 2012.
- [22] D. Brice, M. Marinov, and B. Rüger, "A newly designed baggage transfer system implemented using event-based simulations," *Urban Rail Transit*, vol. 1, no. 4, pp. 194–214, 2015.
- [23] V. Ghilas, E. Demir, and T. V. Woensel, "A scenario-based planning for the pickup and delivery problem with time windows, scheduled lines and stochastic demands," *Transportation Research Part B: Methodological*, vol. 91, pp. 34–51, 2016.
- [24] L. Zhao, H. Li, M. Li et al., "Location selection of intra-city distribution hubs in the metro-integrated logistics system," *Tunnelling and Underground Space Technology*, vol. 80, pp. 246–256, 2018.
- [25] J. Dong, W. Hu, S. Yan, R. Ren, and X. Zhao, "Network planning method for capacitated metro-based underground logistics system," *Advances in Civil Engineering*, vol. 2018, Article ID 6958086, 14 pages, 2018.
- [26] S. Yang, L. Ning, P. Shang, and L. Tong, "Augmented Lagrangian relaxation approach for logistics vehicle routing problem with mixed backhauls and time windows," *Transportation Research Part E: Logistics and Transportation Review*, vol. 135, 2020.
- [27] O. Bräysy, "A reactive variable neighborhood search for the vehicle-routing problem with time windows," *INFORMS Journal on Computing*, vol. 15, no. 4, pp. 347–368, 2003.
- [28] J. de Armas and B. Melián-Batista, "Variable neighborhood search for a dynamic rich vehicle routing problem with time windows," *Computers & Industrial Engineering*, vol. 85, pp. 120–131, 2015.
- [29] A. Dampier and M. Marinov, "A study of the feasibility and potential implementation of metro-based freight transportation in Newcastle upon Tyne," *Urban Rail Transit*, vol. 1, no. 3, pp. 164–182, 2015.
- [30] L. Zhao, X. Wang, J. Stoeter et al., "Path optimization model for intra-city express delivery in combination with subway system and ground transportation," *Sustainability*, vol. 11, no. 3, p. 758, 2019.
- [31] Z. Chen, J. Xue, A. Z. Rose, and K. E. Haynes, "The impact of high-speed rail investment on economic and environmental change in China: a dynamic CGE analysis," *Transportation Research Part A: Policy and Practice*, vol. 92, pp. 232–245, 2016.
- [32] K. Wang, W. Xia, and A. Zhang, "Should China further expand its high-speed rail network? Consider the low-cost carrier factor," *Transportation Research Part A: Policy and Practice*, vol. 100, pp. 105–120, 2017.
- [33] L. Zhao, Y. Zhao, Q. Hu, H. Li, and J. Stoeter, "Evaluation of consolidation center cargo capacity and locations for China railway express," *Transportation Research Part E: Logistics and Transportation Review*, vol. 117, pp. 58–81, 2018.
- [34] Tencent News, *The Most Heart-Warming Metro in History, the Shenzhen Female Train Starts Operation Today*, Tencent News, Shenzhen, China, 2018.
- [35] H. Saeedi, B. Behdani, B. Wiegman, and R. Zuidwijk, "Assessing the technical efficiency of intermodal freight transport chains using a modified network DEA approach," *Transportation Research Part E: Logistics and Transportation Review*, vol. 126, pp. 66–86, 2019.
- [36] W. Sun, Y. Yu, and J. Wang, "Heterogeneous vehicle pickup and delivery problems: formulation and exact solution," *Transportation Research Part E: Logistics and Transportation Review*, vol. 125, pp. 181–202, 2019.
- [37] M. L. Fisher, R. Jaikumar, and L. N. Van Wassenhove, "A multiplier adjustment method for the generalized assignment problem," *Management Science*, vol. 32, no. 9, pp. 1095–1103, 1986.
- [38] D. Z. W. Wang, H. Liu, and W. Y. Szeto, "A novel discrete network design problem formulation and its global optimization solution algorithm," *Transportation Research Part E:*

- Logistics and Transportation Review*, vol. 79, pp. 213–230, 2015.
- [39] J. Muren, J. Wu, L. Zhou, Z. Du, and Y. Lv, “Mixed steepest descent algorithm for the traveling salesman problem and application in air logistics,” *Transportation Research Part E: Logistics and Transportation Review*, vol. 126, pp. 87–102, 2019.
- [40] N. Mladenović and P. Hansen, “Variable neighborhood search,” *Computers Operations Research*, vol. 24, no. 11, pp. 1097–1100, 1997.
- [41] P. Hansen, N. Mladenović, R. Todosijević, and S. Hanafi, “Variable neighborhood search: basics and variants,” *EURO Journal on Computational Optimization*, vol. 5, no. 3, pp. 423–454, 2017.
- [42] M. Dražić, C. Lavor, N. Maculan, and N. Mladenović, “A continuous variable neighborhood search heuristic for finding the three-dimensional structure of a molecule,” *European Journal of Operational Research*, vol. 185, no. 3, pp. 1265–1273, 2008.
- [43] Ningbo Rail Transit, *Ningbo Urban Rapid Rail Transit Construction Plan (2013-2020)*, Ningbo Rail Transit, Ningbo, China, 2013.
- [44] Dada Express, *Business Introduction*, 2020.
- [45] X. Hu, G. Wang, X. Li, Y. Zhang, S. Feng, and A. Yang, “Joint decision model of supplier selection and order allocation for the mass customization of logistics services,” *Transportation Research Part E: Logistics and Transportation Review*, vol. 120, pp. 76–95, 2018.
- [46] M. Xiao, K. Cai, and H. A. Abbass, “Hybridized encoding for evolutionary multi-objective optimization of air traffic network flow: a case study on China,” *Transportation Research Part E: Logistics and Transportation Review*, vol. 115, pp. 35–55, 2018.
- [47] L. Zhen, C. Ma, K. Wang, L. Xiao, and W. Zhang, “Multi-depot multi-trip vehicle routing problem with time windows and release dates,” *Transportation Research Part E: Logistics and Transportation Review*, vol. 135, 2020.
- [48] R. Cheng and M. Gen, *Genetic Algorithms and Engineering Design*, John Wiley, Hoboken, NJ, USA, 1997.
- [49] J. A. Pazour, R. D. Meller, and L. M. Pohl, “A model to design a national high-speed rail network for freight distribution,” *Transportation Research Part A: Policy and Practice*, vol. 44, no. 3, pp. 119–135, 2010.
- [50] M. Bi, S. He, and W. Xu, “Express delivery with high-speed railway: definitely feasible or just a publicity stunt,” *Transportation Research Part A: Policy and Practice*, vol. 120, pp. 165–187, 2019.
- [51] T. D’Alfonso, C. Jiang, and V. Bracaglia, “Would competition between air transport and high-speed rail benefit environment and social welfare?” *Transportation Research Part B: Methodological*, vol. 74, pp. 118–137, 2015.
- [52] S. Jia, C. Zhou, and C. Qin, “No difference in effect of high-speed rail on regional economic growth based on match effect perspective?” *Transportation Research Part A: Policy and Practice*, vol. 106, pp. 144–157, 2017.
- [53] D. Sun and S. Guan, “Measuring vulnerability of urban metro network from line operation perspective,” *Transportation Research Part A: Policy and Practice*, vol. 94, pp. 348–359, 2016.
- [54] G. Nian, F. Chen, Z. Li, Y. Zhu, and D. Sun, “Evaluating the alignment of new metro line considering network vulnerability with passenger ridership,” *Transportmetrica A: Transport Science*, vol. 15, no. 2, pp. 1402–1418, 2019.
- [55] D. Zhang, “High-speed train control system big data analysis based on the fuzzy RDF model and uncertain reasoning,” *International Journal of Computers Communications & Control*, vol. 12, no. 4, pp. 577–591, 2017.



## Retraction

# Retracted: Supply Chain Decision Analysis of Community E-Commerce Platform under Different Power Structures: Considering the Influence of Value Cocreation

### Computational Intelligence and Neuroscience

Received 17 October 2023; Accepted 17 October 2023; Published 18 October 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] Z. Liu and Y. Li, "Supply Chain Decision Analysis of Community E-Commerce Platform under Different Power Structures: Considering the Influence of Value Cocreation," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 2522245, 10 pages, 2021.



## Research Article

# Supply Chain Decision Analysis of Community E-Commerce Platform under Different Power Structures: Considering the Influence of Value Cocreation

Ziyu Liu  and Yaping Li 

*School of Economics and Management, Hebei University of Science and Technology, Shijiazhuang, China*

Correspondence should be addressed to Ziyu Liu; purpleyuliu@163.com

Received 18 October 2021; Revised 10 November 2021; Accepted 16 November 2021; Published 30 November 2021

Academic Editor: Daqing Gong

Copyright © 2021 Ziyu Liu and Yaping Li. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to explore the impact of different decision-making methods on the profits of various entities in the supply chain of the community e-commerce platform, this paper adopts the method of the Stackelberg game. For the community e-commerce platform supply chain composed of suppliers, community e-commerce platforms, and grid station service providers, considering the degree of supplier value cocreation efforts, this paper studies the optimal decisions under centralized decision-making, supplier-led decentralized decision-making, and community e-commerce platform-led decentralized decision-making, respectively. The results show that the supply chain obtains the highest profit in centralized decision-making; under decentralized decision-making, the dominant party will get higher profits; and the supplier value cocreation sensitivity coefficient is positively correlated with sales price, value cocreation effort level, and total supply chain value. The results are helpful to improve the competitiveness of the community e-commerce platform supply chain in the market and are of great significance to the long-term development of the community e-commerce industry.

## 1. Introduction

The community e-commerce platform is a shopping platform that has emerged in recent years, such as Jingxi Pinpin, Meituan Preferred, and Duoduomai, gradually participating in people's lives. Buying goods online and picking up offline, this combination of online and offline shopping provides consumers with many conveniences. Consumers can purchase products using mobile apps, reducing the inconvenience of traditional offline shopping, and meet the fast pace of modern life, which is widely accepted by people. At the same time, there are more and more community e-commerce platforms, and the competition between platforms has become increasingly fierce. Therefore, while providing convenience to consumers, community e-commerce platforms should also pay attention to their own profits. The competition between community e-commerce platforms has gradually escalated to between supply chains, considering value cocreation is an effective way for the supply chain to

gain a competitive advantage. By integrating resources between enterprises and the interaction between enterprises and consumers, corporate profits and supply chain profits can be improved. With the deepening of consumer theory, the value cocreation centered on consumers has become the source of obtaining new competitive advantages [1]. Through consumers' value cocreation, enterprises can obtain information about consumers' needs, increase customer loyalty and supply chain value by providing products or services that satisfy consumers. At the same time, the dominant position in the supply chain is very important, which directly affects its optimal decision and maximum profit.

## 2. Literature Review

Research on supply chain value cocreation is mostly based on case studies to analyze the specific content and impact of value cocreation in a certain field. Wan et al. connected the

open innovation subject with the value creation model of supply chain and conducted numerical analysis on supply chain value and consumer value [2]. Shen et al. found that a large part of the value creation of brand community e-commerce companies is through the positive contributions of consumers [3]. Hein et al. constructed a platform enterprise ecosystem model based on value creation, analyzed its value creation system, and studied the cooperative relationship between all parties, expounding the value relationship between all stakeholders [4].

Previous studies on supply chain decision focused on the impact of different power structures on decision-making. Luo et al. established a multistage game model to study the impact of different power structures on pricing decisions, retailer, and manufacturer profits [5]. Ma et al. believed that only the dominant manufacturer can benefit from the wholesale price strategy, while both manufacturers and retailers can benefit from the channel strategy [6]. Zhao and Li established a Nash game model and a Stackelberg game model for a low-carbon supply chain consisting of manufacturers and retailers [7]. Gao et al. considered sales effort and explored the impact of different rights structures on the optimal decision-making and performance of a closed-loop supply chain [8, 9].

At present, there are few studies on the community e-commerce platform, mainly focusing on its development and model. Liu analyzed the development status and mode of fresh community e-commerce and believed its market demand and potential are huge [10]. Some scholars have also studied consumers' purchasing intention. Li explored the differences in purchasing intention and influencing factors of different consumers in community e-commerce [11]. Jian and Yang constructed a consumer purchase decision model of community e-commerce and explored the mechanism of community e-commerce characteristics and trust on consumers' purchasing intention [12].

The current research on supply chain value cocreation is still in the aspect of the case and empirical research, and there is very little research on model analysis. Therefore, this paper takes the community e-commerce platform supply chain as the research object, and supplier value cocreation effort level is considered. It explores the impact of different power structures on sales prices, sales volume, and profit creation.

### 3. Research Method

**3.1. Stackelberg Game Method.** The Stackelberg game method is the main research method used in this article. It is a game model proposed by economist Stackelberg in the 1930s to reflect asymmetric competition between enterprises. Assuming that there are two participants in the game, enterprise 1 first decides its output, and enterprise 2 makes output decisions based on enterprise 1's decision. Therefore, enterprise 1 must consider how enterprise 2 will respond when making decisions.

Since the decision-making sequence is involved in the game, the position of each player should be judged first. The dominant position is the leader of the game, and the other is

the follower. The follower's reaction to the market is easily predicted by the leader, and the leader makes decisions that are beneficial to him based on the response of the followers. Most of the current research on supply chain decision-making regards the supplier as the leader of the supply chain. According to the retailer's sales forecast and market demand function, the supplier makes the optimal decision combining with the cost and other factors to determine the wholesale price. Then, retailers determine retail prices based on wholesale prices.

Nevertheless, in real life, with the development of online sales channels such as e-commerce and community e-commerce platforms, the position of each participant in the supply chain is changing gradually. Taking the community e-commerce platform supply chain as an example, some community e-commerce platforms may play a leading role in the supply chain. Therefore, the impact of the decision-making approach on the profitability of the supply chain and each participant has to be fully considered. Because community e-commerce platforms and suppliers occupy an important position in the supply chain, grid station service providers are in a noncore position, only responsible for the selection and distribution of goods, and do not involve operation decisions within the supply chain such as product pricing and order quantities. Therefore, the grid station service provider is no longer added to the decision-making of the supply chain. Based on this, there are three decision-making methods: centralized decision-making in the community e-commerce platform supply chain, decentralized decision-making led by the supplier, and decentralized decision-making led by the community e-commerce platform.

### 3.2. Related Theories

**3.2.1. Supply Chain Decision Theory.** According to the decision-making mode, it can be divided into centralized decision-making and decentralized decision-making.

Centralized decision-making refers to the collective decision-making of the main participants in the supply chain and takes the maximization of the overall benefit of the supply chain as the decision-making goal [13]. However, in fact, centralized decision-making is an ideal state, participants often pursue the maximization of their own profits, and it is difficult to achieve the maximization of overall profit.

Decentralized decision-making means that each participant in the supply chain makes decisions independently and pursues their own best profits rather than the maximization of overall profit.

Section 5 discusses the situation of the community e-commerce platform supply chain under centralized and decentralized decision-making. Among them, decentralized decision-making is divided into two situations. One is a supplier-led supply chain, where the community e-commerce platform acts as a follower and makes relevant decisions based on the supplier's strategy. The other is dominated by community e-commerce platforms, where suppliers act as platform followers and make strategies based on platform decisions.

**3.2.2. Market Demand Theory.** Market demand is the total amount of goods that a particular group of consumers may purchase in a given region, time, and circumstances [14]. Moreover, it is affected by total demand and price. Higher prices decrease the quantity demanded, and lower prices increase the quantity demanded, expressed by formula  $Q = a - bP$ , where  $a$  is the total market demand,  $b$  is the sensitivity coefficient of demand to price, and the main factors affecting demand include the product price, consumer preference, and other factors.

In this paper, the basic hypothesis of Section 4 uses the market demand theory to represent the relationship between market demand and cost faced by community e-commerce platforms.

**3.2.3. Diminishing Marginal Effect.** The diminishing marginal effect refers to the continuous input of a certain element under other conditions remaining unchanged [14]. When the input increases to a certain amount, the increment that can be achieved by the unit element becomes less. In other words, the cost per unit increment is going to be higher.

In this paper, the basic hypothesis in Section 4 uses the theory of diminishing marginal effect to represent the cost of suppliers participating in value cocreation.

## 4. Result

**4.1. Model Description.** The SCG supply chain studied in this paper is a three-level supply chain composed of suppliers (S), community e-commerce platform (C), and grid station service provider (G). The structure is shown in Figure 1. The supplier is responsible for the production and supply of goods. The community e-commerce platform provides a platform for consumers to browse and order goods. It sets up goods on the platform and publishes information on goods and activities. Consumers determine the purchase by browsing relevant information. In addition, the community e-commerce platform is also responsible for collecting consumer feedback on goods. Grid station service provider picks and distributes goods according to consumer order content.

This paper considers the value cocreation efforts of suppliers in the decision-making of this three-level supply chain. As the source of the supply chain, suppliers provide goods or services for the entire supply chain and sell them through the community e-commerce platform [15]. The quality of goods or services directly affects consumer experience and satisfaction. Only when consumers approve the goods or services will they repeat purchases of the goods or services in the future, thereby increasing the sales volume of the supply chain and occupying a unique competitive advantage in the fierce market competition. Suppliers' value cocreation behavior is reflected in all aspects of product production and design. In terms of product design, suppliers widely absorb consumer demand and opinions and design and improve goods according to demand to meet more consumers. In terms of production, select more appropriate

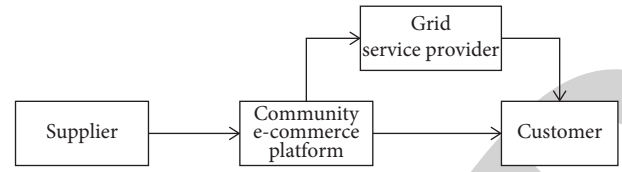


FIGURE 1: Supply chain structure of community e-commerce platform.

raw materials and processes to meet consumer demand for quality and so on.

### 4.2. Basic Assumptions

**Hypothesis 1:** Based on the hypothesis of economic man and rational man, all the research subjects in this paper are completely rational. They aim to maximize their own profits and assume information sharing among enterprises in the chain.

**Hypothesis 2:** Assume that the wholesale price of the supplier to the community e-commerce platform is  $W$  and production cost is  $C_1$ . The selling price of the community e-commerce platform is  $P$ , and the cost of the unit product is  $C_2$ . The income of the grid station service provider for product selection and distribution per unit is  $Z$ .

**Hypothesis 3:** Assume the value cocreation effort cost of the supplier is  $C(v) = (1/2)\alpha v^2$ . According to the diminishing marginal effect, the marginal utility of unit cost is diminishing when the level of effort becomes higher, and the cost of unit effort is increasing. Among them,  $\alpha$  is the supplier's value cocreation cost coefficient,  $v$  is the supplier's value cocreation effort level, and the value range is  $0 \leq v \leq 1$ .

**Hypothesis 4:** Assuming that the market demand is related to the price and the degree of cocreation of supplier value, the market demand function is  $Q = a - bP + sv$ . Among them,  $P$  and  $Q$  are the price and sales volume of the product, respectively,  $a$  represents the maximum capacity of the market,  $b$  represents the sensitivity coefficient of demand to price, and  $s$  is the sensitivity coefficient of supplier value cocreation, where  $a$ ,  $b$ , and  $s$  are constants greater than zero.

**4.3. Model Construction.** According to the above model description and basic assumptions of the SCG community e-commerce platform supply chain, the value function of each participant can be obtained.

The value of the supplier is

$$\pi_1 = (W - C_1)Q - \frac{1}{2}\alpha v^2. \quad (1)$$

The value of the community e-commerce platform is

$$\pi_2 = (P - W - C_2 - Z)Q. \quad (2)$$

The grid station service provider value is

$$\pi_3 = ZQ. \quad (3)$$

The supply chain value of the SCG community e-commerce platform is

$$\pi_{\text{SCG}} = \pi_1 + \pi_2 + \pi_3 = (P - C_1 - C_2)Q - \frac{1}{2}\alpha v^2. \quad (4)$$

## 5. Discussion

**5.1. Centralized Decision-Making.** Under centralized decision-making, all enterprises in the supply chain aim to achieve the maximum profit of the supply chain. Jointly decide the optimal price  $P_c^*$  and supplier value cocreation effort level  $v_c^*$  together. At this point, the profit function of supply chain decision is

$$\pi_{\text{SCG}} = (P - C_1 - C_2)(a - bP + sv) - \frac{1}{2}\alpha v^2. \quad (5)$$

First, calculate the second derivative of  $P$  and  $v$  separately from equation (5) to obtain the Hessian matrix:

$$H = \begin{bmatrix} \frac{\partial^2 \pi_{\text{SCG}}}{\partial P^2} & \frac{\partial^2 \pi_{\text{SCG}}}{\partial P \partial v} \\ \frac{\partial^2 \pi_{\text{SCG}}}{\partial v \partial P} & \frac{\partial^2 \pi_{\text{SCG}}}{\partial v^2} \end{bmatrix} = \begin{bmatrix} -2b & s \\ s & -\alpha \end{bmatrix}. \quad (6)$$

Since  $|H| \geq 0$ , the second-order Hessian matrix is a positive definite matrix, and  $\pi_{\text{SCG}}$  is a strictly concave function of  $P$  and  $v$ . At this time, there is a unique optimal solution.

Combine first-order conditions  $(\partial \pi_{\text{SCG}} / \partial P) = 0$  and  $(\partial \pi_{\text{SCG}} / \partial v) = 0$ , which can be solved as

$$P_c^* = \frac{\alpha a + (ab - s^2)c_1 + c_2}{2ab - s^2}, \quad (7)$$

$$v_c^* = \frac{sa - sb(c_1 + c_2)}{2ab - s^2}.$$

Substituting the result into the market demand function, we can obtain

$$Q_c^* = \frac{\alpha ab - \alpha b^2(c_1 + c_2)}{2ab - s^2}. \quad (8)$$

At this time, the overall maximum profit of the SCG supply chain is

$$\pi_{\text{SCG}} = \frac{\alpha[a - b(c_1 + c_2)]^2}{2(2ab - s^2)}. \quad (9)$$

In summary, we can get Conclusion 1.

**Conclusion 1.** In the case of centralized decision-making, the optimal decision of SCG community e-commerce platform supply chain is  $P_c^* = (\alpha a + (ab - s^2)c_1 + c_2) / (2ab - s^2)$ ,  $v_c^* = (sa - sb(c_1 + c_2)) / (2ab - s^2)$ ,  $Q_c^* = (\alpha ab -$

$\alpha b^2(c_1 + c_2)) / (2ab - s^2)$ , and the maximum profit of the supply chain is  $\pi_{\text{SCG}} = \alpha[a - b(c_1 + c_2)]^2 / 2(2ab - s^2)$ .

## 5.2. Decentralized Decision-Making

**5.2.1. Community E-Commerce Platform-Led Decentralized Decision-Making.** In decentralized decision-making, the members of the supply chain make decisions independently and take their own maximum profit as the decision-making goal. In the case of community e-commerce platform leading, the community e-commerce platform first considers its own maximum profit. The decision sequence is as follows: firstly, the community e-commerce platform decides the sales price  $P$  and the order quantity  $Q$ , then the supplier decides the wholesale price  $W$  and the supplier value cocreation effort level  $v$ , and uses the backward induction method of the Stackelberg game to solve the model [16].

By backward induction, first calculate the response function of the supplier, then substitute the profit function of the community e-commerce platform to obtain the optimal decision of the community e-commerce platform and, finally calculate the optimal value of each subject. Since directly substituting the wholesale price  $W$  into the supplier's profit function cannot calculate the optimal solution, assume that the supplier's profit per unit of a commodity is  $n$ , and the sales price  $P$  is expressed as  $P = W + n$ .

Therefore, the profit function of suppliers in the SCG supply chain is

$$\pi_1 = (W - C_1)Q - \frac{1}{2}\alpha v^2, \quad (10)$$

$$W - C_1[a - b(w + n) + sv] - \frac{1}{2}\alpha v^2.$$

Similar to the derivation process of centralized decision-making, the second derivative of  $W$  and  $v$  can be obtained by calculating the second derivative of  $W$  and  $v$ , respectively, to obtain the Hessian matrix:

$$H_2 = \begin{bmatrix} \frac{\partial^2 \pi_1}{\partial W^2} & \frac{\partial^2 \pi_1}{\partial W \partial v} \\ \frac{\partial^2 \pi_1}{\partial v \partial W} & \frac{\partial^2 \pi_1}{\partial v^2} \end{bmatrix} = \begin{bmatrix} -2b & s \\ s & -\alpha \end{bmatrix}. \quad (11)$$

Since  $|H_2| \geq 0$ , we can get that  $\pi_1$  is a strictly concave function about  $W$  and  $v$ , and  $\pi_1$  has the only optimal solution.

Combine first-order conditions  $(\partial \pi_1 / \partial W) = 0$  and  $(\partial \pi_1 / \partial v) = 0$ , which can be solved as

$$W_1^* = \frac{\alpha a + bc_1 - s^2 c_1 - \alpha bn}{2ab - s^2}, \quad (12)$$

$$v_1^* = \frac{s[a - bc_1 - bn]}{2ab - s^2}.$$

Next, find the optimal strategy of the community e-commerce platform. The profit function of the community e-commerce platform is

$$\pi_2 = (P - W - C_2 - Z)(a - bp + sv). \quad (13)$$

Substitute  $P = W + n$  into the profit function and simplify to obtain

$$\pi_2 = (n - C_2 - Z)(a - bw - bn + sv). \quad (14)$$

Find the second partial derivative of  $n$  with respect to  $\pi_2$ ,  $(\partial^2 \pi_2 / \partial n^2) = -2b < 0$ . Therefore,  $\pi_2$  is a concave function about  $n$ , and there is a unique optimal solution. Letting  $(\partial \pi_2 / \partial n) = 0$ , we can get

$$n_1^* = \frac{a - bc_1 + b(C_2 + Z)}{2b}. \quad (15)$$

Substituting  $n^*$  into  $W^*$  and  $v^*$ , we can get

$$W_1^* = \frac{\alpha[a + b(3c_1 - c_2 - Z)] - 2s^2 c_1}{2(2ab - s^2)}, \quad (16)$$

$$v_1^* = \frac{s[a - b(c_1 + c_2 + Z)]}{2(2ab - s^2)}.$$

Substituting  $n^*$  and  $W^*$  into the function  $P = W + n$ , the sales price  $P^*$  of the community e-commerce platform can be obtained:

$$P_1^* = \frac{ab^2((c_1 + c_2 + Z) + b[3\alpha a - s^2(c_1 + c_2 + Z)]) - as^2}{2b(2ab - s^2)}. \quad (17)$$

Substituting  $P^*$  into the market demand function  $Q = a - bp + sv$ ,  $Q^*$  can be obtained by calculation and simplification:

$$Q_1^* = \frac{ab[a - b(c_1 + c_2 + Z)]}{2(2ab - s^2)}. \quad (18)$$

**Conclusion 2.** Under the decentralized decision-making led by the community e-commerce platform, the optimal strategy of the SCG supply chain is the sales price

$$\begin{aligned} P_1^* &= \frac{ab(c_1 + c_2 + Z) + b[3\alpha a - s^2(c_1 + c_2 + Z)] - as^2}{2b(2ab - s^2)}, \\ Q_1^* &= \frac{ab[a - b(c_1 + c_2 + Z)]}{2(2ab - s^2)}, \\ v_1^* &= \frac{s[a - b(c_1 + c_2 + Z)]}{2(2ab - s^2)}. \end{aligned} \quad (19)$$

**5.2.2. Supplier-Led Decentralized Decision-Making.** In the case of supplier-led, the supplier first considers the maximization of its own interests. The decision sequence is as follows: the supplier decides the wholesale price  $W$  and the supplier value cocreation effort level  $v$ , and then the community e-commerce platform decides the sales price  $P$  and order quantity  $Q$  of the product.

First, calculate the optimal strategy of the community e-commerce platform. The profit function of the community e-commerce platform is

$$\pi_2 = (P - W - C_2 - Z)Q = (P - W - C_2 - Z)(a - bp + sv). \quad (20)$$

Find the second partial derivative of  $P$  with respect to  $\pi_2$ ,  $(\partial^2 \pi_2 / \partial P^2) = -2b$ . Therefore,  $\pi_2$  is a strictly concave function, and there is a unique optimal solution. So letting  $(\partial \pi_{SCG} / \partial P) = 0$ , we can get the optimal sales price of community e-commerce platform  $P_2^*$ :

$$P_2^* = \frac{a + sv + b(w + z + c_2)}{2b}. \quad (21)$$

Next, solve the supplier's optimal strategy, and the supplier's profit function is

$$\pi_1 = (W - C_1)Q = W - C_1 a - bp + sv - \frac{1}{2}\alpha v^2. \quad (22)$$

Substitute  $P_2^*$  into equation (22) and simplify

$$\begin{aligned} \pi_1 &= aw - w^2 b - bcw + svw - ac_1 + wbc_1 \\ &\quad + bc_1(c_2 + z) - svc_1 - \frac{1}{2}\alpha v^2. \end{aligned} \quad (23)$$

Calculating the second derivative of  $W$  and  $v$ , respectively, for equation (23), we can get the Hessian matrix:

$$H_3 = \begin{bmatrix} \frac{\partial^2 \pi_1}{\partial W^2} & \frac{\partial^2 \pi_1}{\partial W \partial v} \\ \frac{\partial^2 \pi_1}{\partial v \partial W} & \frac{\partial^2 \pi_1}{\partial v^2} \end{bmatrix} = \begin{bmatrix} -2b & s \\ s & -\alpha \end{bmatrix}. \quad (24)$$

Since  $|H_3| \geq 0$ , we can get that  $\pi_1$  is a strictly concave function about  $W$  and  $v$  and  $\pi_1$  has the only optimal solution.

Combining the first-order conditions  $(\partial \pi_1 / \partial W) = 0$  and  $(\partial \pi_1 / \partial v) = 0$ , we can solve it

$$W_2^* = \frac{a\alpha + ab(c_1 - Z - c_2) - s^2 c_1}{2ab - s^2}, \quad (25)$$

$$V_2^* = \frac{s[a - b(z + c_1 + c_2)]}{2ab - s^2}.$$

Substituting  $W^*$ ,  $v^*$  into  $P$ ,  $P^*$ ,  $Q^*$  can be calculated

$$\begin{aligned} P_2^* &= \frac{3a\alpha + [(b\alpha - 2s^2)(z + c_1 + c_2)]}{2(2ab - s^2)}, \\ Q_2^* &= \frac{ab[a - b(c_1 + c_2 + Z)]}{2(2ab - s^2)}. \end{aligned} \quad (26)$$

**Conclusion 3.** Under the supplier-led decentralized decision-making, the optimal price is  $P_2^* = (3a\alpha + [(b\alpha - 2s^2)(z + c_1 + c_2)])/2(2ab - s^2)$ , and  $Q_2^* = ab[a - b(c_1 + c_2 + Z)]/2(2ab - s^2)$ .

$$2(2\alpha b - s^2), W_2^* = (\alpha a + ab(c_1 - Z - c_2) - s^2 c_1)/(2\alpha b - s^2), \\ v_2^* = s[a - b(z + c_1 + c_2)]/(2\alpha b - s^2).$$

5.3. *Comparative Analysis of Models.* Through the above solution, the optimal strategy in these three cases can be obtained, as shown in Table 1.

By sorting out the results in Table 1, the optimal sales price, supplier value cocreation effort level, and supplier wholesale price under different decentralized decisions have the following relationship:  $P_1^* < P_2^*$ ,  $W_1^* < W_2^*$ ,  $V_1^* > V_2^*$ .

*Proof.*  $P_1^* - P_2^* = (s^2[b(c_{1+c_2+Z}) - a]/2b(2\alpha b - s^2)) < 0$ ,  $v_1^* - v_2^* = -(s[a - b(z + c_1 + c_2)]/(2\alpha b - s^2)) < 0$ ,  $W_1^* - W_2^* = (s[b(z + c_1 + c_2) - a]/2(2\alpha b - s^2)) < 0$ .

This shows that different dominant modes will bring different optimal strategies to the supply chain of the community e-commerce platform and have different effects on the supply chain. In the case of community e-commerce platform dominance, the optimal selling price, supplier value cocreation effort, and the optimal wholesale price are all smaller than the optimal selling price in the case of supplier dominance.  $\square$

5.4. *Numerical Analysis.* In the calculation above, we have obtained a series of results under different decision modes. In order to further analyze and verify the optimal decisions of community e-commerce platforms, suppliers, and grid service providers under different decision modes, we will conduct numerical analysis in this section.

By referring to previous literature, this paper assumes that the values of each parameter are as follows:  $a=100$ ,  $b=0.7$ ,  $c_1=7$ ,  $c_2=2$ ,  $Z=1$ ,  $\alpha=0.2$ ,  $s=0.15$ . The optimal decision and maximum profit under different decision modes are shown in Table 2.

Change the value of  $s$ , so  $a=100$ ,  $b=0.7$ ,  $c_1=7$ ,  $c_2=2$ ,  $Z=1$ ,  $\alpha=0.2$ ,  $s=0.2$ . The optimal decision and maximum profit under different decision modes are shown in Table 3.

It can be seen from Table 2 that the results calculated by the calculation example are consistent with the conclusions drawn from the above model analysis.

- (1) The first thing that can be obtained is that when the SCG supply chain makes decentralized decision-making, the optimal decision will be affected by the dominant power. Under the supply chain dominated by community e-commerce platforms, the wholesale price of suppliers is 43.12, the sales price is 112.54, and the supplier value cocreation effort level is 27.08. When the supplier leads, the wholesale price is 79.23, the sales price is 118.35,

and the supplier value cocreation effort level is 54.17. When the community e-commerce platform dominates the supply chain, to maximize its own profits, the community e-commerce platform uses the dominant advantage of the supply chain to lower the wholesale price of suppliers as much as possible, while suppliers keep their profits from falling sharply by lowering the level of value cocreation. When the supplier is the leader of the supply chain, it will keep the wholesale price at a high level as much as possible and increase the level of value cocreation to meet the needs of more consumers and increase sales. Through these two methods to obtain their own maximum profits.

- (2) When decentralized decision-making, different power structures will bring different effects to community e-commerce platforms, suppliers, grid station service providers, and even the entire supply chain. When the community e-commerce platform is the leader of the supply chain, the profit of the community e-commerce platform is 1679.1, the profit of the supplier is 839.78, the profit of the grid station service provider is 25.28, and the total profit of the supply chain is 2544.16. When the supplier occupies the leading position in the supply chain, the profit of the community e-commerce platform is 913.11, the profit of the supplier is 1532.53, the profit of the grid station service provider is 25.28, and the total profit of the SCG community e-commerce platform supply chain is 2470.93. It can be seen that the dominant power will have different effects on the profits of different entities in the supply chain. When the community e-commerce platform is the dominant player, the profit of the community e-commerce platform is greater than that of the supplier. When suppliers dominate, supplier profits are much higher than community e-commerce platform profits.
- (3) Under centralized decision-making, the maximum profit of the SCG supply chain is 3409.59. The overall profits of the supply chain under different decentralized decisions are 2544.16 and 2470.93, respectively, which are less than the profits obtained under the centralized decision-making model.
- (4) It can be seen from Tables 2 and 3 that whether it is in centralized decision-making or decentralized decision-making, when the supplier value cocreation sensitivity coefficient  $s$  increases, the sales price  $P$ , supplier value cocreation level  $v$ , and supply chain total value  $\pi$  all increase accordingly. Therefore, it



TABLE 1: Comparison of the supply chain of community e-commerce platforms under different decision-making modes.

Centralized decision	Decentralized decision-making	
	Community e-commerce platform dominates	Supplier-led
$P^*$	$(\alpha a + (\alpha b - s^2)c_1 + c_2)/(2\alpha b - s^2)$	$(3\alpha a + [(b\alpha - 2s^2)(z + c_1 + c_2)]/2(2\alpha b - s^2)) / (2\alpha b - s^2)$
$V^*$	$(sa - sb(c_1 + c_2))/(2\alpha b - s^2)$	$s[a - b(z + c_1 + c_2)] / (2\alpha b - s^2)$
$Q^*$	$(\alpha ab - \alpha b^2(c_1 + c_2))/(2\alpha b - s^2)$	$\alpha b[a - b(c_{1+c_2+z})] / 2(2\alpha b - s^2)$
$W^*$		$(\alpha a + \alpha b(c_1 - c_2 - Z) - s^2 c_1) / (2\alpha b - s^2)$

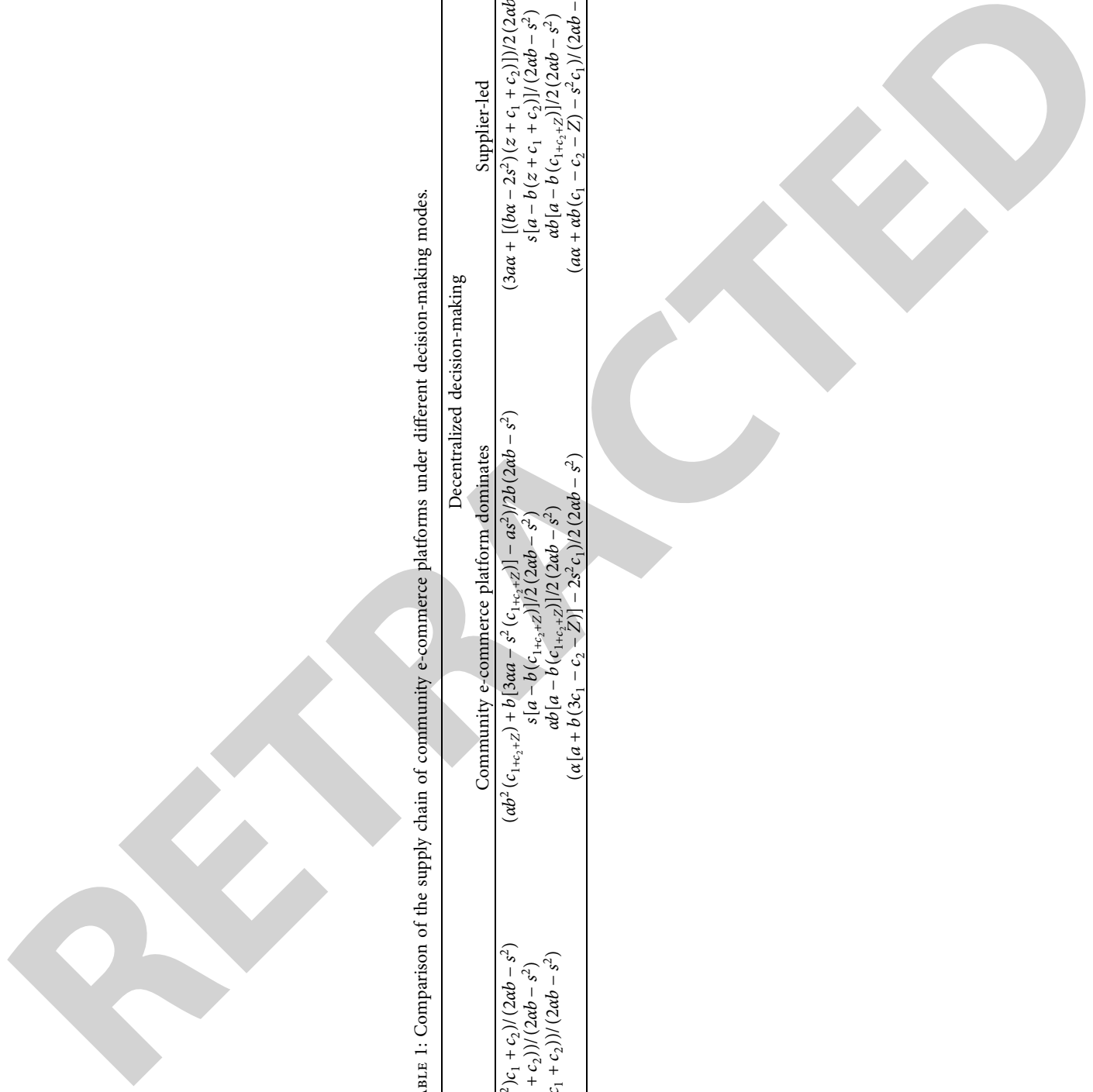


TABLE 2: Results analysis of different decision-making modes of SCG community e-commerce platform supply chain.

Centralized decision	Decentralized decision-making	
	Community e-commerce platform dominates	Supplier-led
$P^*$	81.77	112.54
$\nu^*$	54.58	27.08
$Q^*$	50.94	25.28
$W^*$		43.12
$\pi_1$		839.78
$\pi_2$		1679.1
$\pi_3$		25.28
$\pi_{SCG}$	3409.59	2544.16

TABLE 3: Results analysis of different decision-making modes of SCG community e-commerce platform supply chain.

Centralized decision	Decentralized decision-making	
	Community e-commerce platform dominates	Supplier-led
$P^*$	87.08	115.18
$V^*$	78.08	38.75
$Q^*$	54.66	27.13
$W^*$		45.75
$\pi_1$		901.13
$\pi_2$		1802.24
$\pi_3$		27.13
$\pi_{SCG}$	3658.2	2730.5

can be seen that  $s$  is positively correlated with sales price  $P$ , supplier value cocreation level  $\nu$ , and total supply chain value  $\pi$ .

## 6. Conclusion

This paper is aimed at a community e-commerce platform supply chain composed of suppliers, community e-commerce platforms, and grid station service providers and considers the value cocreation of suppliers. It analyzes the centralized decision-making and decentralized decision-making of the SCG model and gets the following conclusions.

- (1) Different decision-making modes will bring different optimal strategies to the supply chain. The profit generated by the supply chain in centralized decision-making is greater than that generated by decentralized decision-making, and different power structures will bring different revenue situations to enterprises in the supply chain. Therefore, to maintain a high degree of coordination, all companies in the supply chain need to focus on coordination and cooperation to provide a good foundation for centralized decision-making. In order to ensure that companies in the supply chain can obtain information from each other timely and accurately, some emerging information technologies such as blockchain, big data, and Internet of Things technologies can be introduced into the supply chain of community e-commerce platforms. In addition, some measures must be taken to guarantee

cooperation between supply chains. Enterprises in supply chains can find a reasonable profit distribution from the perspective of profit redistribution to benefit all enterprises in supply chains [17].

- (2) Under decentralized decision-making, when the supplier is the leader, the effort cost of the supplier participating in value cocreation is higher than that when the community e-commerce platform is the leader. In this case, the wholesale price of goods will be higher than the price of the community e-commerce platform to obtain higher income. The position of the enterprise in the supply chain affects the profit of the enterprise directly. When it is a leader, the profit reaches the maximum, and when it is a follower, the profit is the smallest. Therefore, companies can improve their competitive advantages, enhance their irreplaceability in the supply chain, and strive for supply chain leadership to obtain their maximum profits. The publicity and sales activities of the platform itself and the products sold on the platform are valued by the community e-commerce platform. Consumers' opinions and suggestions on the platform and products, paying attention to the emotional value of the consumers, are widely absorbed by it [18]. The platform has been satisfied and recognized by consumers by improving the service level from consumers' perspective, and the number of audiences has been increased [19]. In this way, the community e-commerce platform can increase its brand influence, thereby improving its position in the supply chain; they improve their income and contribute to the overall efficiency of the supply chain. Suppliers can consolidate their position in consumers and the supply chain by improving the quality of their products and producing unique products to gain leadership advantages [20].
- (3) Whether in centralized or decentralized decision-making, the sensitivity of supplier value cocreation is positively correlated with community e-commerce platform prices, supplier value cocreation efforts, and total supply chain profits. The greater the sensitivity, the higher the profitability of the supply chain. This shows that the increase in

the supplier value cocreation sensitivity coefficient can increase the profit of the community e-commerce platform, the profit of the supplier, and the overall profit of the supply chain. Therefore, suppliers are more willing to participate in value cocreation and take some measures to increase the sensitivity coefficient, such as cultivating consumers' consumption habits and adopting value cocreation methods that are more acceptable to consumers.

Through combing the literature, it is found that the existing literature mainly studies the model and development trend of the community e-commerce platform supply chain, and there are few studies on the supply chain decision-making. In addition, there are few studies on supply chain value cocreation from the perspective of model analysis. Therefore, the main contribution of this paper is to introduce the degree of supplier value cocreation effort to the decision-making model of community e-commerce platform supply chain and explore the influence of supplier value cocreation on sales price, retail price, and profit of each enterprise, which provides a theoretical basis for the development of community e-commerce platform supply chain. However, this research is not unlimited. Firstly, when establishing the decision-making model, it is assumed that the members of the community e-commerce platform supply chain have already realized information sharing, and the noninteroperability of information is not considered. Secondly, there are often multiple suppliers and multiple commodities. This paper only studies the situation of a single supplier and single commodity. Therefore, in future research, other factors and subjects will be further considered in the model to obtain a conclusion closer to the implementation situation.

### Data Availability

The data used to support the findings of this study are included within the article.

### Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

### Acknowledgments

This paper was financially supported by the 100 Outstanding Innovative Talents Support Project of Hebei Province (SLRC2019005) and the Social Science Foundation Project of Hebei Province (HB20GL011), which made the present work possible.

### References

- [1] S. L. Vargo and R. F. Lusch, "Evolving to a new dominant logic for marketing," *Journal of Marketing*, vol. 68, no. 1, pp. 1–17, 2004.
- [2] X. L. Wan, T. T. Hao, and X. X. Rong, "Supply chain value creation considering open innovation from the perspective of Co-creation," *Chinese Journal of Management Science*, vol. 7, pp. 57–66, 2017.
- [3] X.-L. Shen, Y.-J. Li, Y. Sun, and Y. Zhou, "Person-environment fit, commitment, and customer contribution in online brand community: a nonlinear model," *Journal of Business Research*, vol. 85, no. 85, pp. 117–126, 2018.
- [4] A. Hein, J. Weking, M. Schrieck, M. Wiesche, M. Böhm, and H. Krömer, "Value co-creation practices in business-to-business platform ecosystems," *Electronic Markets*, vol. 29, no. 3, pp. 503–518, 2019.
- [5] Z. Luo, X. Chen, J. Chen, and X. Wang, "Optimal pricing policies for differentiated brands under different supply chain power structures," *European Journal of Operational Research*, vol. 47, no. 6, pp. 345–378, 2016.
- [6] L. Ma, R. Zhang, S. Guo, and B. Liu, "Pricing decisions and strategies selection of dominant manufacturer in dual-channel supply chain," *Economic Modelling*, vol. 29, no. 6, pp. 2558–2565, 2012.
- [7] H. J. Zhao and Z. Li, "Dynamic game analysis of R&D investment in low carbon supply chain," in *Proceedings of the 28th Chinese Control and Decision Conference*, pp. 4570–4573, Yinchuan, China, May 2016.
- [8] J. Gao, H. Han, L. Hou, and H. Wang, "Pricing and effort decisions in a closed-loop supply chain under different channel power structures," *Journal of Cleaner Production*, vol. 112, pp. 2043–2057, 2016.
- [9] M. L. Xu, X. Qi, and H. Y. Jian, "The decision of closed-loop supply chain considering value Co-creation under different authoring structure," *Journal of Industrial Technological Economics*, vol. 39, no. 8, pp. 62–71, 2020.
- [10] Y. Liu, "Research on the operation mode of the community E-commerce supply chain of fresh agricultural products," *Journal of Commercial Economics*, vol. 4, pp. 102–105, 2021.
- [11] L. Y. Li, "Consumer differences in purchase intention and influencing factors of fresh agricultural products in community E-commerce: based on empirical analysis of 578 consumers in nanchang city," *Journal of Agro-Forestry Economics and Management*, vol. 19, no. 4, pp. 457–463, 2020.
- [12] H. Y. Jian and H. Yang, "The impact of community e-commerce technical characteristics on consumers' purchase intention," *Journal of Industrial Technological Economics*, vol. 23, pp. 99–102, 2020.
- [13] J. Su, C. Li, Q. Zeng, J. Yang, and J. Zhang, "A green closed-loop supply chain coordination mechanism based on third-party recycling," *Sustainability*, vol. 11, no. 19, Article ID 103390, 2019.
- [14] H. Y. Gao, *Western Economics*, People's Publishing House, Beijing, China, 2007.
- [15] T. Maiti and B. C. Giri, "A closed loop supply chain under retail price and product quality dependent demand," *Journal of Manufacturing Systems*, vol. 37, pp. 624–637, 2015.
- [16] B. Niladri, "Syam and amit. Pazgal , " Co-creation with production externalities," *Marketing Science*, vol. 32, no. 5, pp. 805–820, 2013.
- [17] J. Y. Chen, "Research on the realization mechanism of multi-value Co-creation in the retail industry under the O2O business ecosystem: taking yonghui supermarket as an example," *Journal of University of Science and Technology Beijing*, vol. 37, no. 1, pp. 74–83, 2021.
- [18] T. C. Zhang, M. F. Jahromi, and M. Kizildag, "Value cocreation in a sharing economy: the end of price wars?" *International Journal of Hospitality Management*, vol. 71, pp. 51–58, 2018.
- [19] H. Kaur, M. Paruthi, J. Islam, and L. Hollebeek, "The role of brand community identification and reward on consumer brand engagement and brand loyalty in virtual brand

## Retraction

# Retracted: A Novel Hybrid Algorithm for Multiobjective Location-Allocation Problem in Emergency Logistics

### Computational Intelligence and Neuroscience

Received 1 August 2023; Accepted 1 August 2023; Published 2 August 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] H. Chu and Y. Chen, "A Novel Hybrid Algorithm for Multiobjective Location-Allocation Problem in Emergency Logistics," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 1951161, 12 pages, 2021.

## Research Article

# A Novel Hybrid Algorithm for Multiobjective Location-Allocation Problem in Emergency Logistics

Hongrui Chu <sup>1</sup> and Yahong Chen<sup>2</sup>

<sup>1</sup>*School of Management and Engineering, Capital University of Economics and Business, Beijing 100070, China*

<sup>2</sup>*Beijing Intelligent Logistics System Collaborative Innovation Center, School of Information, Beijing Wuzi University, Beijing 101149, China*

Correspondence should be addressed to Hongrui Chu; [chuhongrui@cueb.edu.cn](mailto:chuhongrui@cueb.edu.cn)

Received 20 October 2021; Accepted 15 November 2021; Published 28 November 2021

Academic Editor: Daqing Gong

Copyright © 2021 Hongrui Chu and Yahong Chen. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Increased frequency of disasters keeps reminding us of the importance of effective resource distribution in postdisaster. To reduce the suffering of victims, this paper focuses on how to establish an effective emergency logistics system. We first propose a multiobjective optimization model in which the location and allocation decisions are made for a three-level logistics network. Three objectives, deprivation costs, unsatisfied demand costs, and logistics cost, are adopted in the proposed optimization model. Several cardinality and flow balance constraints are considered simultaneously. Then, we design a novel effective IFA-GA algorithm by combining the firefly algorithm and genetic algorithm to solve this complex model effectively. Furthermore, three schemes are proposed to improve the effectiveness of the IFA-GA algorithm. Finally, the numerical results provide several insights on the theory and practice of relief distribution, which also illustrate the validity of the proposed solution algorithm.

## 1. Introduction

Large-scale natural or man-made disasters, such as the 2010 Haitian Earthquake, the 2011 tsunami in Japan, the civil war in Syria, and the 2019 coronavirus disease, have occurred frequently throughout the world, resulting in tremendous consequences of enormous casualties and property losses. The world has witnessed a steadily increased number in both disasters and affected people since the 1900s. After large-scale disasters, a great need for medical and daily supplies will be invoked in the affected areas, which lack adequate relief resources. Thus, the rapid distribution of external commodities is critical to mitigating the losses caused by disasters. Therefore, emergency logistics has received wide attention from both practical managers and scholars [1]. However, how to design an efficient emergency logistics system is a big challenge [2]. Before the last decade, most research on emergency management focused on the responses of public servants, government agencies, and insurance firms in times of crisis [3]. Emergency logistics, also

commonly known as humanitarian logistics, is emerging and becoming a hot spot of operations management in the recent decade.

In most of the current related research, emergency logistics problems are studied based on the analytical formulations of the traditional commercial sector, unlike the commercial scenario, whose decision objective is to minimize the logistics cost. In emergency relief cases, the scientific community is encouraged to provide effective logistics systems to reduce human suffering. Therefore, the adapted objectives in commercial and emergency logistics are radically different. The modified analytical formulations of the traditional logistics are not very suitable for emergency logistics problems [4].

To derive an appropriate objective function for emergency logistics, Holguín-Veras et al. [5] incorporated welfare economic principles into emergency management and introduced the deprivation costs, which express the economic value of human suffering caused by a lack of access to resources or services. Since then, some studies using

deprivation costs have been conducted in the emergency logistics field [6]. For example, Pérez-Rodríguez and Holguín-Veras [7] developed an inventory-allocation-routing model for the optimal critical supplies by minimizing logistics and social costs. Ismail [3] studied a relief distribution problem by considering the logistics and deprivation costs. Some other papers also considered human suffering in the emergency logistics field. Chapman and Mitchell [8] studied the distribution centers selection problem by minimizing the suffering of the population. Cotes and Cantillo [9] developed a facility location model for repositioning supplies in preparation for disasters. Their formulation considers deprivation costs in the objective function. Rivera-Royero et al. [10] proposed a dynamic model to serve rescue demand. They considered the level of urgency of demand points. Moreno et al. [11] presented a novel model to optimize location, transportation, and fleet sizing for emergency logistics with considering deprivation costs. Cantillo et al. [12] assessed transportation network vulnerability from logistical costs and deprivation costs. Zhu et al. [13] studied the emergency relief routing optimization problem considering injured degree and deprivation cost.

Similar to the above literature, we adopt the deprivation costs to characterize human suffering in this paper. We study a location-allocation problem of emergency logistics. Three objectives, deprivation costs, unsatisfied demand losses, and logistics costs, are adopted in the proposed optimization model. The fairness objective of this paper equalizes the supplies between affected areas, which has the same purpose as unmet demand minimization in [14]. In addition, our literature also lies in the stream of multiobjective programming, which is widely used to solve emergency optimization problems. Sun et al. [15] proposed a robust optimization model to decide the facility location, resource allocation, and casualty transportation in a three-level rescue chain composed of casualty clusters, temporary facilities, and general hospitals. Their objectives are to minimize the total Injury Severity Score (ISS) and logistics cost. Zhou et al. [2] studied the multiperiod dynamic emergency resource scheduling problem by a proposed multiobjective optimization model. Cao et al. [16] proposed a multiobjective nonlinear programming model for resource distribution regarding the beneficiary perspective on sustainability. The model is solved by maximizing the lowest victims' perceived satisfaction and minimizing the largest deviation of victims' perceived satisfaction. Çelik et al. [17] studied the stochastic debris clearance problem among postdisaster operations. They determined a sequence of roads to clear with the goal of maximizing satisfied relief demand. Gralla et al. [18] developed a method to value the objective functions of emergency logistics based on expert preferences over five key attributes. They found that the amount of cargo delivered is the most valued objective and cost the least important. In a similar vein, we conduct the optimization problem based on a three-level emergency logistics network. Considering the complexity of reality, we formulate the problem by a multiobjective mixed-integer nonlinear programming with cardinality and flow balance constraints.

Most optimization problems of emergency logistics are hard to solve effectively by exact algorithms, especially in the limited available time. In such situations, adopting a metaheuristic algorithm is imperative [19]. The third related literature stream of this paper is algorithm design. Zhou et al. [2] designed an evolutionary algorithm based on decomposition (MOEA) to solve a multiobjective multiperiod dynamic emergency resource scheduling problem. Cao et al. [16] adopted genetic algorithm to solve a relief distribution model. Haghgi et al. [20] proposed a multiobjective location and transportation programming model, which is solved by a nondominant sorting genetic algorithm. Eisenhandler and Tzur [21] modeled a food rescue problem as a routing resource allocation problem with the equitable allocations objective. They presented a heuristic approach based on the large neighborhood search framework to solve the model. Zhang et al. [22] determined the truck and drone routes in a humanitarian relief network by combining column generation and tabu search algorithms. This paper designs a metaheuristic algorithm named IFA-GA by combining the firefly algorithm and genetic algorithm. Some improved schemes are embedded into IFA-GA to obtain better solutions.

The contributions of this paper can be summarized as follows: first, we propose a new multiobjective location-allocation optimization model for emergency logistics, in which deprivation costs, fairness, and operation costs are considered. The numerical results show that deprivation costs have a significant effect on location decisions. Second, we design a new IFA-GA algorithm based on the combination of the firefly algorithm and genetic algorithm. Three schemes, population initialization, search strategy, and update mechanism of population, are proposed to improve IFA-GA. The computational results illustrate that the proposed IFA-GA has better solving performance than the other algorithms.

The remainder of this paper is organized as follows: Section 2 presents some presumptions and notations. The detailed descriptions of the three objective functions and model formulation are also given in Section 2. Section 3 develops the hybrid IFA-GA algorithm for solving the proposed model and gives the improved schemes. Section 4 conducts numerical studies and analyses. Finally, Section 5 presents conclusions and future work.

## 2. Problem Description and Formulation

In this section, we first give some presumptions and notations. Then, three relief-related objective functions are proposed to represent human suffering, fairness, and logistics cost. Finally, we present a multiobjective location and resource allocation model for three-level emergency logistics.

*2.1. Presumptions and Notations.* Emergency supplies mainly come from two sources: local rescue warehouses or external dispatching. Local rescue warehouses can provide supplies immediately after disasters occur, but the amount is



limited. External dispatching mainly comes from the Strategic National Stockpile (SNS), which integrates different parties' resources and provides the primary resources to the affected areas by railway or airplane in the subsequent relief phase. In this paper, we consider an emergency logistics problem in the second case. Most emergency logistics systems have a three-tier structure, including SNS, regional transfer center (RTC), and local distribution point (LDP). In reality rescue practice, large amounts of emergency resources must be transported from SNS to affected points as soon as possible. To improve the efficiency of the rescue, RTCs are proposed as hubs to bridge SNS and LDP. This work proposes a joint facility location and resource allocation model for this three-level emergency logistics system. Some assumptions used in our model are as follows:

- (1) There are one SNS, some RTCs, and a fixed number of LDPs. The resources are first transported from SNS to RTCs and then transported from RTCs to LDPs. LDPs are equivalent to affected points that have resource demands. SNS, RTCs, and LDPs have enough capacity to store resources. The capacity constraints are ignored in our optimization model.
- (2) The demand quantity of resources in each affected point can be predicted, and the total available quantity of resources is also known.
- (3) The rescue resources are transported by trucks. Moreover, there are enough available trucks in each supply point.
- (4) The locations of RTCs are decided in the postdisaster phase. After establishing RTCs, all roads from RTCs to SNS and LDPs are fixed. This work does not consider the uncertainty of road networks. Therefore, the delivery time between two supply points is determined by the speed of the truck.
- (5) The rescue demand has an additional time-related attribute. If the actual delivery time exceeds the required threshold, the deprivation costs of victims will be invoked to represent the economic valuation of the human suffering associated with a lack of access to a good or service.

The RTCs are established in selected reasonable sites the first time after the disaster occurs. In commercial logistics, RTCs can quickly move goods from the central warehouse, manufacturers to retailers. In emergency logistics, large-scale disaster is characterized by a wide range of affected points, great demand, and long duration. Therefore, the appropriate location of RTCs has a more critical role in improving the transport efficiency of resources. With the above assumptions, the location decision of RTCs, the allocation decisions from SNS to RTCs and from RTCs to LDPs will be studied in our proposed multiobjective optimization model. The notations used in the following text are listed as follows:

- (i)  $O$ : the SNS point;
- (ii)  $I$ : the candidate location set of RTCs,  $i = 1, 2, \dots, I$ ;

- (iii)  $J$ : the location set of LDPs,  $j = 1, 2, \dots, J$ ;
- (iv)  $d_{Oi}$ : the distance from SNS  $O$  to RTC candidate  $i$ ;
- (v)  $d_{ij}$ : the distance from RTC candidate  $i$  to LDP  $j$ ;
- (vi)  $C_{OI}$ : the unit transport cost from SNS to RTC;
- (vii)  $C_{IJ}$ : the unit transport cost from RTC to LDP;
- (viii)  $C_i$ : the construction cost of RTC  $i$ ;
- (ix)  $CV_i$ : the unit operation cost of RTC  $i$ ,  $CV_i$  multiplied by inventory equal to the total operation cost of RTC  $i$ ;
- (x)  $D_j$ : the demand quantity of LDP  $j$ ;
- (xi)  $Q_o$ : the initially available quantity of the resource in SNS;
- (xii)  $Q_i$ : the initially available quantity of the resource in RTC  $i$ ;
- (xiii)  $\pi_{jt}$ : the unit deprivation cost of each victim in LDP  $j$  at time  $t$ ;
- (xiv)  $\delta_j$ : the deprivation costs of LDP  $j$  at time  $t$ ;
- (xv)  $l_{oj}$ : the loss function of unsatisfied demand of LDP  $j$ ;
- (xvi) OC: the total operation cost of emergency logistics;
- (xvii)  $t_{ij1}$ : the first delivery time of resources from RTC  $i$  to LDP  $j$ ;
- (xviii)  $t_{ij2}$ : the second delivery time of resources from RTC  $i$  to LDP  $j$ ;
- (xix)  $v$ : the speed of truck;
- (xx)  $z_i$ : 0–1 variables, if candidate  $i$  is selected as RTC,  $z_i = 1$ , else,  $z_i = 0$ ;
- (xxi)  $x_{oi}$ : the allocated quantity of emergency resources from SNS to RTC  $i$ ;
- (xxii)  $y_{ij1}$ : the first allocated quantity of emergency resources from RTC  $i$  to LDP  $j$ ;
- (xxiii)  $y_{ij2}$ : the second allocated quantity of emergency resources from RTC  $i$  to LDP  $j$ ;
- (xxiv)  $\varepsilon_{ij}$ : auxiliary variable,  $\varepsilon_{ij} = 1$  if RTC  $i$  provides resources to LDP  $j$ , else  $\varepsilon_{ij} = 0$ .

**2.2. Multiobjective Optimization Model.** There are three objectives in our model that reflect three service levels of emergency logistics: (1) minimizing the total deprivation costs of victims, (2) minimizing the loss of unsatisfied demand, and (3) minimizing the total cost of the emergency logistics system. The formulations consider a relief group that first delivers critical supplies from SNS to RTCs and then delivers goods from RTCs to demand nodes. Furthermore, RTCs also have a limited initial inventory of critical supplies. Therefore, there are twice transportations from RTC  $i$  to LDP  $j$  if RTC  $i$  is selected as a hub to supply LDP  $j$ . RTC  $i$  first provides resources by its own inventory. The first resources reach LDP  $j$  at time  $t_{ij1}$ . Then, RTC  $i$  transfers the second resources which come from SNS and arrive at LDP  $j$  at time  $t_{ij2}$ . During a planning horizon  $T$ , the relief group's decisions are which RTCs should be selected,

how much, and when to deliver to the demand nodes in need of supplies.

**2.2.1. The Deprivation Costs.** Human suffering is one of the critical factors in humanitarian logistics. The theory of deprivation costs [5] provides an excellent method to bring human suffering factors into the decision objective. Conceptually, deprivation cost is the economic value of human suffering caused by a lack of access to resources or services.

$$\delta_j = \pi(t_{ij1})\min(D_j, y_{ij1}) + \pi(t_{ij2})\min(\max(D_j - y_{ij1}, 0), y_{ij2}) + \pi(T)\max(D_j - y_{ij1} - y_{ij2}, 0), \quad (1)$$

where  $t_{ij1}$  is dependent on the distance between RTC  $i$  and LDP  $j$ .  $t_{ij2}$  is dependent on the distance from SNS to LDP  $j$  via RTC  $i$ . The first part of  $\delta_j$  is the deprivation costs of the satisfied demand in the first supply. The second part of  $\delta_j$  is the deprivation costs of the satisfied demand in the second supply. The third part of  $\delta_j$  is the deprivation costs of the remaining unsatisfied demand at the end.

It should be noted that the beneficiaries had already experienced a deprivation cost at time  $t < t_{ij1}$ . Therefore, the

$$\sum_{j \in J} \delta_j = \sum_{i \in I, j \in J} at_{ij1}^2 \min(D_j, y_{ij1}) + at_{ij2}^2 \min(\max(D_j - y_{ij1}, 0), y_{ij2}) + aT^2 \max(D_j - y_{ij1} - y_{ij2}, 0). \quad (2)$$

**2.2.2. The Loss of Unsatisfied Demand.** In the context of emergency rescue, decision-makers should try their best to meet the supplies needs in each affected point, but the goods are inadequate at the first time. Therefore, another important objective in emergency logistics is to keep fairness. In this work, we provide a loss function of unsatisfied demand to evaluate the fairness of rescue. The loss function  $l_{oj}$  is an S type curve of unsatisfied demand in LDP  $j$ . This reflects that the unsatisfied demand has a learning effect on the emergency service level. If the unsatisfied demand takes the maximum value, the loss function becomes constant, and no more losses will occur. In addition, developed regions tend to be densely populated, which have more severe consequences if emergency demands are not met. Thus, the demand quantity is considered in the loss function, which can be given by

$$l_{oj} = \max(D_j - y_{ij1} - y_{ij2}, 0)P(h_j), \quad (3)$$

where  $P(h_j) = be^{-h_j/(1-h_j)}$  and  $h_j = (\sum_{i \in I} y_{ij1} + y_{ij2}/D_j)$ .  $h_j$  is the proportion of satisfied demand in LDP  $j$ . Obviously,  $1 - h_j$  is the ratio of unsatisfied demand.  $P(h_j)$  is related to service level, which converges to loss upper bound  $b$  if  $h_j$  tends to 0, and 0 if  $h_j$  tends to 1. Thus, we use the above loss function to quantify the fairness in this paper. The total unsatisfied demand loss of emergency logistics is given as follows:

Suppose LDPs at which identical individuals experience the same unit deprivation cost  $\pi(t)$  at time  $t$ . According to the theory of deprivation costs, the deprivation costs function  $\delta_j$  is discontinuous, increases with deprivation time, and then drops to zero when a delivery large enough to fulfill the needs of all individuals is received. The total deprivation costs at node  $j$  are equal to the aggregation of individual costs:

classical deprivation costs theory evaluates the cumulative cost function up to time  $T$ . For simplification, this work only considers the deprivation costs at the arrival times of supplies and the end time of the period. The unite deprivation cost  $\pi(t)$  is increasing convex function with respect to  $t$ ,  $0 < t \leq T$ . In this paper, we use a quadratic function  $\pi(t) = at^2$  to specify deprivation cost and  $a$  is a cost coefficient. Thus, we can easily obtain the total deprivation costs at the end of period  $T$ , which can be expressed as

$$\sum_{j \in J} l_{oj} = \sum_{i \in I, j \in J} \max(D_j - y_{ij1} - y_{ij2}, 0)P(h_j). \quad (4)$$

**2.2.3. The Operation Cost.** The operation cost of emergency logistics includes three parts: the construction cost of RTCs, the transport cost, and the inventory holding cost in RTCs. In the setting of this paper, the SNS already exists before disasters occur. The LDPs will distribute the relief resources as soon as they are received. Thus, we only consider the holding cost of RTCs because they are newly built. Then, the operation cost objective can be given by

$$OC = \sum_{i \in I} [z_i C_i + C_{oi} d_{oi} x_{oi} + \sum_{j \in J} C_{ij} d_{ij} (y_{ij1} + y_{ij2}) + CV_i (x_{oi} + Q_i)]. \quad (5)$$

**2.3. Mathematical Formulation.** Based on the above analysis, we suppose that the emergency manager's goals are to minimize the deprivation costs, the loss of unsatisfied demand, and the operation cost. Without loss of generality, we adopt two weight parameters  $\alpha$  and  $\beta$  ( $0 < \alpha, \beta, 1 - \alpha - \beta < 1$ ) to simplify this multiobjective optimization to scalar objective optimization. Meanwhile, the manager also needs to consider some realistic constraints, such as path constraint and capacity constraints. Then, the logistics optimization model is formulated as follows:

$$\min \alpha \sum_{j \in J} \delta_j + \beta \sum_{j \in J} lo_j + (1 - \alpha - \beta)OC, \quad (6)$$

$$\text{s.t. } \delta_j = \sum_{i \in I} at_{ij1}^2 \min(D_j, y_{ij1}) + at_{ij2}^2 \min(\max(D_j - y_{ij1}, 0), y_{ij2}) + aT^2 \max(D_j - y_{ij1} - y_{ij2}, 0), \quad \forall j \in J, \quad (7)$$

$$lo_j = \max(D_j - y_{ij1} - y_{ij2}, 0)P(h_j), \quad \forall j \in J, \quad (8)$$

$$P(h_j) = be^{(-h_j/1-h_j)}, \quad \forall j \in J, \quad (9)$$

$$h_j = \frac{\sum_{i \in I} y_{ij1} + y_{ij2}}{D_j}, \quad \forall j \in J, \quad (10)$$

$$OC = \sum_{i \in I} \left[ z_i C_i + C_{oi} d_{oi} x_{oi} + \sum_{j \in J} C_{ij} d_{ij} (y_{ij1} + y_{ij2}) + CV_i (x_{oi} + Q_i) \right], \quad (11)$$

$$t_{ij1} = \frac{d_{ij}}{v}, \quad \forall i \in I, j \in J, \quad (12)$$

$$t_{ij2} = \frac{d_{ij} + d_{oi}}{v}, \quad \forall i \in I, j \in J, \quad (13)$$

$$\sum_{i \in I} \epsilon_{ij} = 1, \quad \forall j \in J, \quad (14)$$

$$\sum_{j \in J} \epsilon_{ij} y_{ij1} = Q_i, \quad \forall i \in I, \quad (15)$$

$$\sum_{j \in J} \epsilon_{ij} y_{ij2} = x_{oi}, \quad \forall i \in I, \quad (16)$$

$$y_{ij1} \leq M \epsilon_{ij} z_i, \quad \forall i \in I, j \in J, \quad (17)$$

$$y_{ij2} \leq M \epsilon_{ij} z_i, \quad \forall i \in I, j \in J, \quad (18)$$

$$\epsilon_{ij} \leq z_i, \quad \forall i \in I, j \in J, \quad (19)$$

$$\sum_{i \in I, j \in J} \epsilon_{ij} (y_{ij1} + y_{ij2}) = \min \left( Q_o + \sum_{i \in I} Q_i, \sum_{j \in J} D_j \right), \quad (20)$$

$$x_{oi} \leq M z_i, \quad \forall i \in I, \quad (21)$$

$$\sum_{i \in I} x_{oi} = Q_o, \quad (22)$$

$$z_i, \epsilon_{ij} \in \{0, 1\}, \quad \forall i \in I, j \in J, \quad (23)$$

$$x_{oi}, y_{ij1}, y_{ij2} \in R^+, \quad \forall i \in I, j \in J. \quad (24)$$

The objective function shown in equation (6) has been described detailedly in the above. Constraints (7), (8), and (11) show the computational formulas of the three objectives, respectively. Constraints (9) and (10) describe the loss

function and the satisfied demand rate used to calculate the second objective. Constraints (11) and (12) describe the relationship between transport distance and delivery time. Constraint (14) restricts that each LDP can only receive

goods from one RTC. Constraints (15) and (16) balance the flows of resources in each RTC for the first and second deliveries. Constraints (17) and (18) guarantee that RTC  $i$  can provide relief supplies to LDP  $j$  only if  $i$  has been selected and  $j$  is allocated to it. Here,  $M$  is a sufficiently large number. Constraint (19) ensures that only a selected RTC can provide emergency supplies to LDPs. Constraint (20) indicates that all goods need to be shipped if the supply is less than demand; else, the delivered quantity is equal to demand. Constraint (21) ensures that only the selected RTC can receive goods from SNS. Constraint (22) indicates the flow balance of resources in SNS. Finally, constraints (23) and (24) define the decision variables.

### 3. A Hybrid Heuristic Solution Approach

The proposed model is one of the location-allocation type problems, most of which is NP-hard. Therefore, the optimal solution of the proposed model is difficult to be obtained in a short time. However, the actual emergency decision-making often has higher requirements on timeliness, which inspires us to adopt metaheuristic algorithms. This section combines an improved firefly algorithm with a genetic algorithm to design a new hybrid algorithm called IFA-GA. Furthermore, some schemes are also proposed to increase the solution quality.

**3.1. The Firefly Algorithm (FA).** FA is a nature-inspired swarm-based optimization algorithm proposed by Yang [23], which is designed to imitate the natural phenomenon of fireflies. FA supposes that all fireflies are homogeneous. Each firefly will be attracted and move close to a brighter firefly. If there are no brighter fireflies nearby, the firefly will move randomly. For an optimization problem, fireflies represent feasible solutions. The brightness of the firefly has a positive relationship with its objective value.

Similar to some other metaheuristic algorithms, FA also successfully implements the exploration and exploitation steps to solve optimization problems. The exploration step is carried out by the movement of fireflies from the current position to a brighter position. The exploitation step is realized by the random movements of the rest fireflies. Suppose  $X_n$  is a solution (firefly) in the  $n$ -th iteration of FA, the update movement equation of this firefly is

$$X_{n+1} = X_n + \beta_0 e^{-\gamma r^2} (\bar{X}_n - X_n) + \alpha_n \varepsilon_n, \quad (25)$$

where  $\beta_0$  is the maximum attractiveness between fireflies.  $r$  is the Euclidean distance between two fireflies  $\bar{X}_n$  and  $X_n$ .  $\gamma$  is the light absorption coefficient that adjusts the change of attraction.  $\alpha_n$  is a weight coefficient between  $[0, 1]$ , and  $\varepsilon_n$  is a random vector in  $[0, 1]$ .

**3.2. The Genetic Algorithm (GA).** GA is initially motivated by the Darwinian principle of evolution through natural selection and genetic mechanisms, which has been successfully applied to a wide range of real-world problems of significant complexity. Unlike the iterative update mechanism in other

algorithms, GA derives a new solution by operating the existing solutions directly. There are two genetic operators in GA: crossover and mutation. A typical design for a classical GA can be given as follows:

- (1) Initialize a population in which chromosomes are generated randomly.
- (2) Calculate the fitness of each chromosome in the population.
- (3) Generate new chromosomes by the following steps: (a) select two chromosomes using the proportional fitness selection, and then cross the selected chromosomes to obtain a child chromosome; (b) select one chromosome using the proportional fitness selection, and apply uniform mutation to it with a mutation rate to produce another child chromosome.
- (4) Update the population. Add new chromosomes to the population and dispose of the worst chromosomes to maintain the size of the population.
- (5) If the stopping criteria have not been met, return to Step 2.

GA is first proposed by Holland [24] to solve computationally intractable problems. The development and success of GAs have greatly enriched the computational approaches for the optimization problem. Particularly, GA is very suitable for combinatorial optimization problems whose structure and character can avoid the tedious coding and decoding works of chromosomes.

**3.3. The IFA-GA Algorithm.** FA has a good convergence speed due to the attraction of fireflies. However, the solution of FA may fall into a local optimum because of its limited exploration capability, especially when the fireflies encounter extreme values. In contrast, GA can explore the solution space widely due to the crossover and mutation operators. However, the convergence rate of GA is relatively slow due to the lack of directionality. Fortunately, both FA and GA have specific and understandable procedures which are easy to implement. Furthermore, both of them are suitable to be hybridized with other optimizers due to their multimodalities. Inspired by these privileges, we introduce an IFA-GA algorithm by combining the desirable characteristics of FA and GA. The emergency logistics model of this paper includes location decisions and allocation decisions. Considering the characteristics of FA and GA algorithms, we plan to use FA to solve location decisions and use GA to solve allocation decisions. Before presenting the detailed implementation of the proposed hybrid algorithm, we give some schemes for improving the algorithm efficiency.

**3.4. Schemes for Improving the IFA-GA Algorithm.** Most of the metaheuristic algorithms would suffer from two problems: metaheuristic algorithms usually have an earlier convergence, and the initial solution and the diversity of the solution space often affect their search quality. To overcome



these two shortcomings, we enhance the proposed IF-GA algorithm through three schemes in this section.

**3.4.1. Scheme 1: Population Initialization.** The initial population has a significant effect on the convergence speed and the quality of the final solutions. In most metaheuristic algorithms, a population is generated by random initialization. In this paper, we initialize the population by a myopic strategy.

Suppose RTCs can be located at any point of two-dimensional space with the cardinality constraint that the total number of RTCs is no more than  $I$ . First, the LDPs are clustered into  $I$  groups by considering the weight of demand. The locations of RTCs are selected at the center of gravity in each group. Then, we get a solution  $\vec{z} = [z_1, z_2, \dots, z_I]$ . The population of location decisions is generated by random initialization around  $\vec{z}$ . Second, for each location decision, allocating the resources according to the proportion of demand, we can obtain an even solution  $\vec{x} = [x_{o1}, x_{o2}, \dots, x_{oI}]$ ,  $\vec{y}_1 = (y_{ij1})$ , and  $\vec{y}_2 = (y_{ij2})$ . Select two elements of  $\vec{x}$  randomly. The first element subtracts a random number which is added to the second element. The random number is not bigger than the first element. The corresponding rows of  $\vec{y}_2$  should be changed accordingly. For  $\vec{y}_1$ , we introduce a random matrix  $\vec{w} = (w_{ij})$ , each row of  $\vec{w}$  only has two nonzero elements in random columns with a random value  $\tilde{w}$  and  $-\tilde{w}$ . Adding  $\vec{y}_1$  to  $\vec{w}$ , we obtain a new  $\vec{y}_1$ . Finally, the initial population is obtained by repeating the above process.

**3.4.2. Scheme 2: The Improved Search Strategy of FA.** In this paper, FA is used to solve the facility location decision of RTCs. As is mentioned above, each decision variable  $z_i$  has a corresponding position. To be more realistic, we do not constrain the positions and only consider a cardinality constraint of RTCs. Therefore, we use coordinate variable  $X_i = (p_i^x, p_i^y)$  to replace  $z_i$  for simplification and define  $X = [X_1, X_2, \dots, X_I]$ . The objective function determines the fitness of firefly  $X$ . The direction and brightness of fireflies depend on the Euclidean distance between them and their fitness.

A dimmer firefly (worse solution) will be attracted by a brighter one (better solution) in FA. This mechanism is beneficial to the convergence of the algorithm. Thus, FA can find the optimal solution quickly. However, FA may fall into premature convergence if it encounters local extrema because the incomplete exploration of nondominated fireflies is one of the main reasons to cause the local convergence of FA. To overcome this weakness, we introduce a new update function for nondominated fireflies, described as follows:

$$X_{n+1} = X_n + \alpha_n(\varepsilon_n - 0.5)X_M, \quad (26)$$

where  $X_M = X_{\text{Max}} - X_n$  if  $\varepsilon_n > 0.5$  and  $X_M = X_n - X_{\text{Min}}$  if  $\varepsilon_n < 0.5$ .  $X_{\text{Max}}$  is the upper bound of search space, and  $X_{\text{Min}}$  is the lower bound. Furthermore, for dominated fireflies, we consider the attraction of the best solution, and the iterative equation is modified as follows:

$$X_{n+1} = X_n + \beta_0 e^{-\gamma r^2} (\tilde{X}_n - X_n) + \beta_0 e^{-\gamma r_{\text{opt}}^2} (\tilde{X}_{\text{opt}} - X_n) + \alpha_n \varepsilon_n, \quad (27)$$

where  $\tilde{X}_{\text{opt}}$  is the brightest firefly in population and  $r_{\text{opt}}$  is the distance between  $\tilde{X}_n$  and  $\tilde{X}_{\text{opt}}$ .

**3.4.3. Scheme 3: The Update Mechanism of Population.** A good algorithm should have an excellent performance in exploring the solution space and have a good ability in convergence. In this paper, the IFA is nested into GA for solving the location problem. GA already has an elimination mechanism that replaces poor solutions to maintain the size of the population. However, the number of eliminated solutions is usually fixed in the original GA. Reference [25] proposed an adaptive mechanism to replace poor solutions dynamically. The number of eliminated solutions decreases in the iterative process. Therefore, the algorithm first tends to explore the solution space better and then moves to the local search. With the same purpose, we introduce the mutated number of GA as follows:

$$q = \left\lceil \text{rand}(0, 1) * \text{PS} * \left(1 - \frac{t}{G_{\text{MAX}}}\right) \right\rceil, \quad (28)$$

where  $\lceil \cdot \rceil$  is the ceiling operator, PS is the population size,  $t$  is the current number of iterations, and  $G_{\text{MAX}}$  is the maximum number of iterations.

The procedure of IF-GA is described in Algorithm 1 in detail.

## 4. Numerical Studies

In order to illustrate the proposed model and validate the performance of IFA-GA, a small-scale instance with one SNS, 10 RTCs, and 40 LDPs is first used. Although the number of nodes in emergency logistics is not considerable, the proposed model has 820 ( $2 * 10 * 40 + 10 + 10$ ) decision variables. It is an extremely complex optimization problem. The results with different objective functions are presented to show the decision differences related to deprivation costs. Furthermore, a comparison is made between IFA-GA, FA, GA, and particle swarm optimization (PSO) based on the same population initialization. The adjustment method for the infeasible solutions in Algorithm 1 is also used in each algorithm to satisfy the constraints  $t$ . Finally, instances extracted from large-scale disasters are used to demonstrate the effectiveness of the designed IFA-GA algorithm furtherly. All experiments are implemented using MATLAB 2019 on a ThinkPad T460s laptop with Intel i5-6200u CPU, 2.3 GHz frequency, and 12 GB memory.

**4.1. Parameter Setting.** In this experiment, we initialize the affected areas in a  $100 * 100$  two-dimensional space randomly. The demand of each affected area is sampled uniformly in  $[50, 150]$ . We assume that the RTCs can be located at any point of the space; then, the distance between a selected RTC and the other node is easily calculated by the

**Input:**

The demands of affected areas. The location and inventory of SNS. The locations and inventories of candidate RTCs. The parameters of the deprivation cost, unsatisfied loss, and operation cost. The weight parameters in the objective function. The parameters of FA and GA algorithms.

**Output:**

The obtained optimal solution set.

**Initialization:**

Initialize the solution population as described in scheme 1. Calculate the distances between points. Calculate the delivery time of each LDP. Then, calculate the values of the fitness function associated with each solution by equation (6).

**Update:**

IFA: represent solutions by fireflies and fix the allocation variables of fireflies. Update the location variables as follows:

- (1) Fireflies are sorted to find the optimal individuals whose  $\text{rand} = 1$  are the nondominated fireflies, and the other individuals are the dominated fireflies.
- (2) For the nondominated fireflies, the update formulations are described as equation (26). For the dominated fireflies, the individuals move following the formulation equation (27). If there are multiple optimal solutions,  $\bar{X}_{\text{opt}}$  in equation (27) is chosen randomly.

GA: recalculating the value of the individual fitness function, the sum of them is denoted by  $\text{SUM}_F$ . Fix the location variables and update the allocation variables as follows:

- (1) Obtain the selection probability based on  $(\text{individual fitness}/\text{SUM}_F)$ , and select two individuals according to this probability. Get the crossover point by an integer at interval  $[1, I + 2J]$ . Do the crossover operation for selected individuals and get two new individuals. Correct the unfeasible individual by elimination and recrossover until all individuals are feasible. The number of newly generated individuals is determined by equation (28).
- (2) Calculate the number of mutated solutions by equation (28). For each mutated solution, select two adjacent variables  $y(i)$  and  $y(i + 1)$  randomly, and execute mutation by  $y(i) + (y_{\text{max}} - ((y(i) + y(i + 1))/2)) * \text{rand}$  and  $y(i + 1) - (y_{\text{max}} - ((y(i) + y(i + 1))/2)) * \text{rand}$ . Correct the unfeasible individual by elimination and remutation until all individuals are feasible.
- (3) An intermediate population is generated by merging the current population with the previous population. Then, the last  $q$  poor individuals in the previous population and the last  $q$  poor individuals in the newly generated population are eliminated from the next generation.

Repeat IFA and GA until the termination criterion is satisfied.

ALGORITHM 1: IFA-GA.

Euclidean distance equation. The cost parameters and other variables in the proposed model are shown in Table 1.

It is essential to select suitable parameters for meta-heuristic algorithms. Additionally, to evaluate the performance of the proposed IFA-GA, we compare it with the standard FA, GA, and PSO. The parameter settings of these algorithms are displayed in Table 2.

#### 4.2. Computational Results regarding Deprivation Costs.

As discussed above, the deprivation cost is an essential factor for the relief decision. To visually illustrate its impact on decision-making, we analyze the location results of RTCs which are shown in Figure 1. Because each RTC is linked to SNS, which is located at point  $(0, 0)$  in our setting, we ignore the paths between SNS and RTCs for simplification. Figure 1 has six pictures. The first three pictures (a), (b), and (c) present the RTCs' locations with a single objective. The subsequent pictures present the RTCs' locations with two and three objectives. Obj1 refers to the deprivation cost objective. Obj2 refers to the unsatisfied demand loss. Obj3 refers to the operation cost. By comparing Figures 1(a) and 1(c), we can find that the locations of RTCs are closer to the affected areas in (a) than in (b). It is realistic because closer to demand will shorten the delivery time. In Figure 1(b), the locations are strange at first glance. It should be noted that objective 2 does not have any path-related cost, so the locations move randomly in the algorithm and have no practical meaning.

TABLE 1: The values of parameters.

Notation	Value
$O$	1
$I$	10
$J$	40
$v$	20
$a$	2
$b$	100
$C_{OI}$	0.08
$C_{IJ}$	0.1
$C_i$	1000
$CV_i$	0.5
$Q_o$	2000
$Q_i$	100

Figures 1(d) and 1(e) confirm the random location results in (b), which have the same results as Figures 1(a) and 1(c), respectively. Figure 1(f) shows the tradeoff location results by objective 1 and objective 3. Figure 1 illustrates that the deprivation cost impacts RTCs' locations, which makes RTCs closer to LDPs.

4.3. Computational Results of Different Algorithms. In this section, we further present several cases to illustrate the effectiveness of the proposed algorithm. For the small-scale case, the iteration results of IFA-GA, FA, GA, and PSD are shown in



TABLE 2: The parameter settings of algorithms.

Algorithm	Parameter	Value
IFA-GA	Population size	200
	Number of iterations	300
	Maximum attractiveness $\beta_0$	0.1
	Absorption parameter $\gamma$	0.001
	Crossover probability	0.5
	Mutation probability	0.2
FA	Population size	200
	Number of iterations	300
	Maximum attractiveness $\beta_0$	0.1
	Absorption parameter $\gamma$	0.001
GA	Population size	200
	Number of iterations	300
	Crossover probability	0.5
	Mutation probability	0.2
PSO	Population size	200
	Number of iterations	300
	Social acceleration coefficient	2
	Personal acceleration coefficient	2

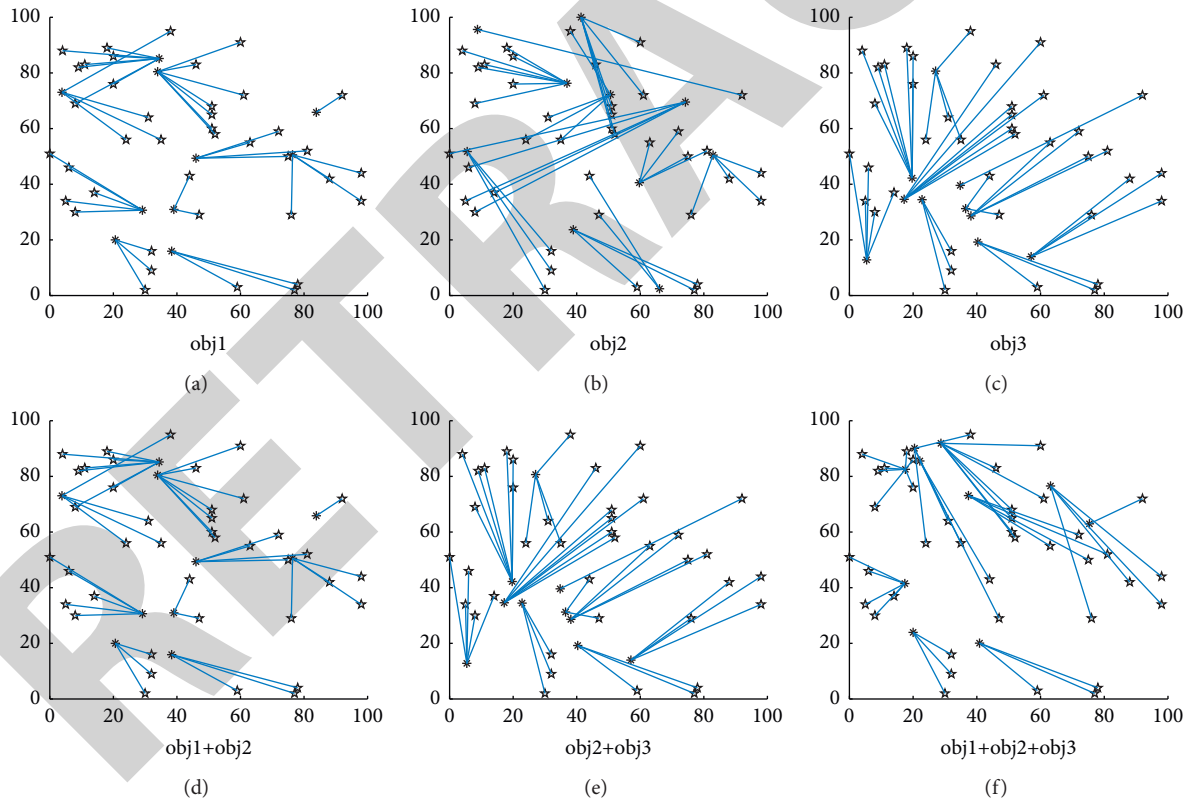


FIGURE 1: The comparison of location decision with the different objective function. (a) Obj1. (b) Obj2. (c) Obj3. (d) Obj1 + Obj2. (e) Obj2 + Obj3. (f) Obj1 + Obj2 + Obj3.

Figure 2. From Figure 2, we can see that the proposed IFA-GA has the best performance in both convergence and objective value. FA and PSD are almost unchanged because the flow balance constraints in the optimization model make the most newly generated solutions infeasible. GA has a certain convergence, but the objective values are much bigger than those of

FA. It should be noted that FA, GA, and PSD are implemented with the same population initialization as IFA-GA. If the population are initialized randomly, the results of FA, GA, and PSD will be much worse.

To compare the results detailedly, we carry out each algorithm 30 times. More detailed results are given in Table 3.

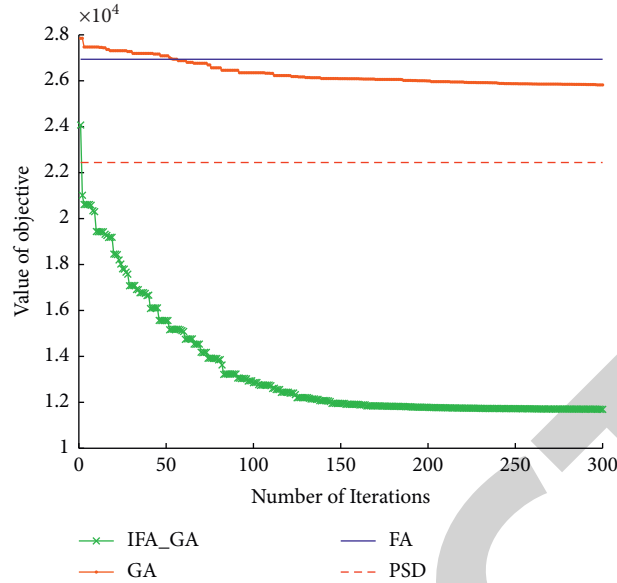


FIGURE 2: The comparison of objective values obtained by different algorithms.

TABLE 3: Numerical results on the small-scale instance.

Algorithm	Mean	Max	Min	STD	Time
IFA-GA	11641	11780	11529	91	51
FA	26246	26972	25586	555	409
GA	23685	23480	23994	191	123
PSO	22468	22242	22838	225	413

From Table 3, we can find that the proposed IFA-GA has better performance than the other algorithms. The mean, max, and min values of the objective obtained by IFA-GA are less than in the other algorithms. The standard deviation of objective values in IFA-GA is much lower than in other algorithms. Furthermore, the run time of IFA-GA is also less than other algorithms. The results of Table 3 indicate that the IFA-GA has better stability and efficiency in solving our proposed model.

Additionally, we conduct experiments with different problem sizes. Four large-scale instances are experimented for further comparison. The results are shown in Table 4.

From Table 4, we can find that the proposed IFA-GA obtains solutions with the best performance than the other algorithms in each instance. In both four cases, IFA-GA obtains the minimum objective value and the lowest deviation. Furthermore, IFA-GA also has the shortest run time. These gaps are to increase along with the problem size increases. FA has the worst performance in run time and stability. GA outdoes FA in each case, with almost the same objective values as PSD but a much lower run time. Additionally, the increase in LDPs brings more objective costs, but the increase in RTCs takes lower objective costs. The result can be obtained by comparing the case  $I = 20, J = 100$  with the case  $I = 50, J = 100$ . The reason is that the increase in RTCs can reduce the deprivation costs effectively, which

TABLE 4: Numerical results under different problem sizes.

Problem size		IFA-GA	FA	GA	PSO
$I = 20$ $J = 50$	Mean	17935	31047	28452	27439
	Max	18240	32567	29497	27883
	Min	17620	30021	27111	27037
	STD	264	1083	889	363
	Time	73	1193	159	607
$I = 20$ $J = 100$	Mean	124675	209152	190568	190039
	Max	125200	211813	192749	191352
	Min	124310	204755	187603	188398
	STD	352	2703	1874	1178
	Time	91	1487	198	756
$I = 50$ $J = 100$	Mean	78051	130882	119106	118934
	Max	79890	135157	122993	122101
	Min	75788	127140	115843	115521
	STD	1467	3048	2574	2338
	Time	139	2071	315	1087
$I = 50$ $J = 150$	Mean	179260	302016	290526	278969
	Max	181800	310912	296233	283487
	Min	177400	291990	287331	274882
	STD	1600	6886	4250	3687
	Time	215	3470	498	1876

enlightens us that more RTCs are recommended to be established when faced with large-scale disasters.

Based on the simulation results, we can derive some practical implications. First, RTCs should be located closer to rescue demand to decrease the human suffering caused by deprivation. Second, in a large-scale disaster scenario, more RTCs should be established to provide fast relief. In the theoretical aspect, we validated the efficiency of the proposed hybrid algorithm, which has an excellent performance in convergence and computation time.

## 5. Conclusions

Different disasters have affected most areas of the world and caused serious casualties and property losses in recent years. An effective emergency logistics occupies a pivotal position to relieve human suffering and mitigate losses after disasters. This paper proposes a multiobjective optimization model for relief distribution. The relief supplies are transferred in a three-level logistics network which includes an SNS, some RTCs, and LDPs. The location decision of RTCs and resource allocation decision from SNS to RTCs and from RTCs to LDPs should be made in postdisasters. We present three objectives, deprivation costs, unsatisfied demand costs, and logistics cost, in this location-allocation optimization model, in which several cardinality and flow balance constraints are considered simultaneously. Realizing the complexity of the problem, we design a novel IFA-GA algorithm by combining classical FA and GA. Three schemes are proposed to improve the effectiveness of the IFA-GA algorithm. The results from numerical studies provide several insights on the theory and practice of relief distribution, which also illustrate the validity of the proposed solution algorithm.

This study is not exhaustive. We only present a simplified model and ignore some factors which will make the model more complex. For example, the uncertainty factors, such as delivery time and quantity, are not considered in the proposed model. Multiple transport and multiple periods are also not investigated in our model. Therefore, we may extend the proposed model to an uncertain situation by using robust optimization theory or distributionally robust theory for future work. Furthermore, more hybrid algorithms that combine exact and metaheuristic approaches should be exploited to solve the proposed optimization model effectively in the future.

## Data Availability

The data were included in the manuscript.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Acknowledgments

This research was supported by the National Natural Science Foundation of China under Grant 72101165, the Fundamental Research Funds for Universities Affiliated to Beijing of Capital University of Economics and Business under

Grant XRZ2020015, and Beijing Intelligent Logistics System Collaborative Innovation Center Foundation under Grant BILSCIEC-2019KF-17.

## References

- [1] R. Z. Farahani, M. M. Lotfi, A. Baghaian, R. Ruiz, and S. Rezapour, "Mass casualty management in disaster scene: a systematic review of OR&MS research in humanitarian operations," *European Journal of Operational Research*, vol. 287, no. 3, pp. 787–819, 2020.
- [2] Y. Zhou, J. Liu, Y. Zhang, and X. Gan, "A multi-objective evolutionary algorithm for multi-period dynamic emergency resource scheduling problems," *Transportation Research Part E: Logistics and Transportation Review*, vol. 99, pp. 77–95, 2017.
- [3] I. Ismail, "A possibilistic mathematical programming model to control the flow of relief commodities in humanitarian supply chains," *Computers & Industrial Engineering*, vol. 157, Article ID 107305, 2021.
- [4] J. Holguín-Veras, M. Jaller, L. N. Van Wassenhove, N. Pérez, and T. Wachtendorf, "On the unique features of post-disaster humanitarian logistics," *Journal of Operations Management*, vol. 30, no. 7–8, pp. 494–506, 2012.
- [5] J. Holguín-Veras, N. Pérez, M. Jaller, L. N. Van Wassenhove, and F. Aros-Vera, "On the appropriate objective function for post-disaster humanitarian logistics models," *Journal of Operations Management*, vol. 31, no. 5, pp. 262–280, 2013.
- [6] J. Shao, X. Wang, C. Liang, and J. Holguín-Veras, "Research progress on deprivation costs in humanitarian logistics," *International Journal of Disaster Risk Reduction*, vol. 42, Article ID 101343, 2020.
- [7] N. Pérez-Rodríguez and J. Holguín-Veras, "Inventory-allocation distribution models for postdisaster humanitarian logistics with explicit consideration of deprivation costs," *Transportation Science*, vol. 50, no. 4, pp. 1261–1285, 2016.
- [8] A. G. Chapman and J. E. Mitchell, "A fair division approach to humanitarian logistics inspired by conditional value-at-risk," *Annals of Operations Research*, vol. 262, no. 1, pp. 133–151, 2018.
- [9] N. Cotes and V. Cantillo, "Including deprivation costs in facility location models for humanitarian relief logistics," *Socio-Economic Planning Sciences*, vol. 65, pp. 89–100, 2019.
- [10] D. Rivera-Royero, G. Galindo, and R. Yie-Pinedo, "A dynamic model for disaster response considering prioritized demand points," *Socio-Economic Planning Sciences*, vol. 55, pp. 59–75, 2016.
- [11] A. Moreno, D. Alem, D. Ferreira, and A. Clark, "An effective two-stage stochastic multi-trip location-transportation model with social concerns in relief supply chains," *European Journal of Operational Research*, vol. 269, no. 3, pp. 1050–1071, 2018.
- [12] V. Cantillo, L. F. Macea, and M. Jaller, "Assessing vulnerability of transportation networks for disaster response operations," *Networks and Spatial Economics*, vol. 19, no. 1, pp. 243–273, 2019.
- [13] L. Zhu, Y. Gong, Y. Xu, and J. Gu, "Emergency relief routing models for injured victims considering equity and priority," *Annals of Operations Research*, vol. 283, no. 1–2, pp. 1573–1606, 2019.
- [14] Y. Liu, H. Lei, Z. Wu, and D. Zhang, "A robust model predictive control approach for post-disaster relief distribution," *Computers & Industrial Engineering*, vol. 135, pp. 1253–1270, 2019.

## *Retraction*

# **Retracted: Intelligent Control of Agricultural Irrigation through Water Demand Prediction Based on Artificial Neural Network**

## **Computational Intelligence and Neuroscience**

Received 17 October 2023; Accepted 17 October 2023; Published 18 October 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

## **References**

- [1] Q. Bo and W. Cheng, "Intelligent Control of Agricultural Irrigation through Water Demand Prediction Based on Artificial Neural Network," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 7414949, 10 pages, 2021.

## Research Article

# Intelligent Control of Agricultural Irrigation through Water Demand Prediction Based on Artificial Neural Network

Qiuyu Bo and Wuqun Cheng 

*Institute of Urban and Rural Construction, Agricultural University of Hebei, Baoding 071001, China*

Correspondence should be addressed to Wuqun Cheng; [chengwuqun@126.com](mailto:chengwuqun@126.com)

Received 9 October 2021; Revised 3 November 2021; Accepted 5 November 2021; Published 23 November 2021

Academic Editor: Daqing Gong

Copyright © 2021 Qiuyu Bo and Wuqun Cheng. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In irrigated areas, the intelligent management and scientific decision-making of agricultural irrigation are premised on the accurate estimation of the ecological water demand for different crops under different spatiotemporal conditions. However, the existing estimation methods are blind, slow, or inaccurate, compared with the index values of the water demand collected in real time from irrigated areas. To solve the problem, this paper innovatively introduces the spatiotemporal features of ecological water demand to the forecast of future water demand by integrating an artificial neural network (ANN) for water demand prediction with the prediction indices of water demand. Firstly, the ecological water demand for agricultural irrigation of crops was calculated, and a radial basis function neural network (RBFNN) was constructed for predicting the water demand of agricultural irrigation. On this basis, an intelligent control strategy was presented for agricultural irrigation based on water demand prediction. The structure of the intelligent control system was fully clarified, and the main program was designed in detail. The proposed model was proved effective through experiments.

## 1. Introduction

As a big agricultural country, China faces a severe shortage of agricultural irrigation water, owing to the low per-capita water resources and the imbalance between water supply and demand [1, 2]. Modern agriculture aims to achieve a high yield of high-quality crops in an ideal cycle. To promote agricultural development in China, it is very meaningful to reconstruct the irrigation water facilities in large irrigated areas, trying to detect an agricultural situation in real time, make reasonable plans for irrigation water consumption, and realize automatic water supply [3–5]. In irrigated areas, the intelligent management and scientific decision-making of agricultural irrigation is premised on the accurate estimation of the ecological water demand for different crops under different spatiotemporal conditions [6–8].

After analyzing the water supply and consumption features in local irrigated areas, Sun et al. [9] adjusted the weights of the indices in the current water consumption evaluation system through analytic hierarchy process (AHP) and constructed a time series prediction model to forecast

the available water resources, industrial water demand, domestic water demand, and agricultural water demand in the irrigated areas with different assurance rates, thereby balancing the water supply with water demand in these areas. Essaid and Caldwell [10] numerically simulated the irrigation water demand of sweet potatoes and peanuts, predicted the accurate water-saving potential of irrigated areas, and put forward an adjustment strategy of the irrigation system to improve the water utilization rate.

To make full use of irrigation water, most irrigated areas choose to renovate the traditional irrigation management mode based on high-resolution remote sensing images. The water demand of irrigated areas is usually inverted and predicted through accurate detection of block changes [11, 12]. Nejatijahromi et al. [13] established a deep semantic segmentation network to classify the planting areas of different types of crops and introduced multiple indices to image fusion during the preprocessing of remote sensing images. In this way, the spectral complexity and spatial resolution were improved, and the feature information of irrigated area images was enhanced. Pani and Mishra [14]



replaced the ordinary convolution and rectified linear unit (ReLU) of the U-Net semantic segmentation network with depthwise separable convolution and Mish activation function, respectively, thereby improving the classification accuracy of the planting areas of different crops. In addition, they constructed an inversion model for the water content in the soil at different depths, which couples the normalized difference vegetation index (NDVI) and automatic water extraction index. Foglia et al. [15] set up a dynamic multifactor prediction model for soil water content. Based on long short-term memory (LSTM) neural network, their model boasts a high prediction accuracy and a low degree of discretization. To predict water demand more accurately, Hassani and Shahdany [16] improved the multiplication of the Jacobian matrix in the Levenberg–Marquardt (LM) algorithm: this operation was replaced with gradient vector calculation and quasi-Hessian matrix. The improved LM algorithm was used as the learning algorithm for the fuzzy neural network of water demand prediction. The results show that the improved algorithm brings higher network learning efficiency and occupies a small storage space. Mimicking the information transmission mode of the cerebral cortex, Li et al. [17] designed a peaking mechanism for the structural growth and pruning of neural network and implemented it to predict the short-term water demand.

To a certain extent, the traditional estimation methods for the water demand in irrigated areas, namely, empirical formulas, hydrological methods, and microclimate methods, are blind and slow. Infrared remote sensing techniques can diagnose the water level by the physiological properties of crops, providing support to decision-makers in water development and management. However, these techniques are not accurate enough. To solve this problem, this paper combines an ANN for water demand prediction with the prediction indices of water demand. Section 2 computes the ecological water demand for nonfixed crops in the agricultural irrigated area, using the simple and direct area quota method. Section 3 predicts the water demand of crops in the agricultural irrigated area by integrating the radial basis function neural network (RBFNN) with autoregression model. Section 4 presents an intelligent control strategy for agricultural irrigation based on water demand prediction and constructs a precision agricultural irrigation intelligent control system, which encompasses a distributed wireless sensor subsystem, a water demand prediction subsystem, and an agricultural irrigation control subsystem. The proposed model was proved effective through experiments.

## 2. Calculation of Ecological Water Demand

The water demand for agricultural irrigation refers to the amount of irrigation water that must be consumed to maintain the normal growth of crops. Given the uneven distribution of population and farmland in China, the planting area of water-demanding regular rice varieties is growing, and the homogeneity of crops is on the rise. To optimize water allocation and realize intelligent control of agricultural irrigation, it is important to scientifically calculate and accurately predict water demand indices of agricultural irrigation.

Firstly, this paper uses the relatively simple and direct area quota method to compute the irrigation water demand of an irrigated area: multiply the size of the irrigated area with the ecological water demand quota of each crop and superimpose the products. Let  $ZS$  be the total water demand of the irrigated area and  $ZS_i$ ,  $P_i$ , and  $Q_i$  the ecological water demand, planting area, and ecological water demand quota of crop  $i$ , respectively. Among them, the quota can be obtained through experimental analysis and theoretical calculation. Assuming that there are  $m$  crops, then  $ZS$  can be calculated by

$$ZS = \sum_{i=1}^m ZS_i = \sum_{i=1}^m P_i \cdot Q_i. \quad (1)$$

In arid areas, crop growth mainly relies on groundwater and irrigation. This paper directly estimates the actual ecological water demand of crops in arid and semiarid areas, using phreatic water evaporation, which requires only a few parameters. Let  $WZS$  be the total ecological water demand of vegetation;  $ZS_{ij}$  and  $DV_{ij}$  the ecological water demand and phreatic water evaporation of crop  $i$  in month  $j$ , respectively; and  $\Psi_V$  the vegetation index of crops. Then,  $ZS$  can be calculated by

$$ZS = \sum_{i=1, j=1}^m ZS_{ij} = \sum_{i=1, j=1}^m P_i \cdot DV_{ij} \cdot \Psi_V. \quad (2)$$

Let  $e_1$  and  $e_2$  be empirically coefficients;  $DB_i$  the buried depth of groundwater at the location of crop  $i$ ;  $DB_{\max}$  the ultimate buried depth for phreatic water evaporation; and  $DV_j^\theta$  the surface evaporation of regular evaporation dish in month  $j$ . Then,  $DV_{ij}$  can be calculated by Aver'yanov's phreatic evaporation formula:

$$DV_{ij} = e_1 \left( \frac{1 - DB_i}{DB_{\max}} \right)^{e_2} DV_j^\theta. \quad (3)$$

During crop growth, transpiration consumes the largest amount of water. In this paper, the actual evapotranspiration is used to determine the ecological water demand of each crop. The calculation assumes that the water supply is sufficient. Let  $DV_j^*$  be the evapotranspiration caused by the transpiration of the reference crop in month  $j$  and  $\Psi_T$  the soil moisture correction coefficient of the irrigated area. Then, the maximum ecological water demand of a crop can be calculated by

$$ZS = \sum_{i=1, j=1}^m ZS_{ij} = \sum_{i=1, j=1}^m P_i \cdot DV_{ij}, \quad (4)$$

where  $DV_{ij}$  can be calculated by

$$DV_{ij} = DV_j^* \times \Psi_V \times \Psi_T. \quad (5)$$

The Penman–Monteith equation, recommended by The United Nations Food and Agriculture Organization (FAO), was adopted to compute the potential evaporation  $DV^*$  of crops in an irrigated area. Let  $\Gamma$  be the slope between the temperature change curve and the saturated vapor pressure;  $FS$  the net radiation of crop surface;  $TH$  the heat flux of the



soil in the irrigated area;  $\delta$  the psychrometer constant;  $T_{AV}$  the mean temperature;  $v_f$  the wind speed at a fixed height; and  $AV_B$  and  $AV_R$  saturated and actual vapor pressures, respectively. Then,  $DV^*$  can be calculated by

$$DV^* = \frac{0.4\Gamma(FS - TH) + \delta(900/(T_{AV} + 273))v_f(AV_B - AV_R)}{\Gamma + \delta(1 + 0.35v_2)}. \quad (6)$$

The slope  $\Gamma$  can be calculated by

$$\Gamma = \frac{4100 \times 0.62(17.3T_{AV}/(T_{AV} + 327.3))}{(T_{AV} + 327.3)^2}. \quad (7)$$

The net radiation of crop surface  $FS$  is the difference between net short-wave radiation  $FS_D$  and net long-wave radiation  $FS_C$ :

$$FS = FS_D - FS_C. \quad (8)$$

Let  $FS_S$  be solar radiation;  $m_S$  the actual sunshine hours; and  $M_S$  the most probable sunshine hours. Then,  $FS_D$  can be calculated by

$$FS_D = 0.78 \left( \frac{1}{4} + \frac{m_S}{2M_S} \right) FS_S. \quad (9)$$

$$FS_C = 4.903 \times 10^{-9} \left( 1.35 \frac{0.25 + 0.5(m_S/M_S)}{0.75 + 2 \times (c/100000)} - 0.35 \right) (0.34 - 0.14\sqrt{AV_R}) \left( \frac{(MT_{\max-j})^4 + (MT_{\min-j})^4}{2} \right). \quad (14)$$

Let  $DT_{\max-k}$  and  $DT_{\min-k}$  be the daily mean maximum and minimum temperatures of day  $k$ ;  $SD_{AV}$  the daily mean relative humidity; and  $v_c$  the wind speed at the altitude of  $c$  meters. Then, the saturated vapor pressure  $AV_B$  can be calculated by

$$AV_B = \frac{r(DT_{\max-k}) + r(DT_{\min-k})}{2}, \quad (15)$$

where  $r$  is exponentiation:

$$r(t) = 0.62e^{17.28t/(t+327.3)}. \quad (16)$$

The actual vapor pressure  $AV_R$  can be calculated by

$$AV_R = \frac{AV_B \times SD_{AV}}{100}. \quad (17)$$

Suppose the height is fixed at 2 m. Then, the mean wind speed can be calculated by

$$v_f = v_c \frac{4.9}{\ln(67.8c - 5.4)}. \quad (18)$$

### 3. RBFNN-Based Water Demand Prediction

**3.1. Network Construction.** To predict the crop water demand in the agricultural irrigated area, this paper sets up a prediction model based on RBFNN, which boasts a strong

approximation ability and a fast convergence speed. As shown in Figure 1, the RBFNN consists of three layers: an input layer that connects and transmits external signals, a hidden layer that realizes RBF mapping, and an output layer that performs linear summation. The input layer receives the sample data about the parameters required to compute the crop water demand in the target irrigated area. The RBF mapping, which underpins the nonlinear transform between the input layer and hidden layer, constitutes the basis of the hidden layer. Therefore, the output of the RBFNN can be regarded as a linear combination of basis functions derived from hidden layer nodes.

Let  $\Phi_S$  be the solar constant;  $\theta$  the latitude; and  $DO$  the day order. Then, solar radiation  $FS_S$  can be calculated by

$$FS_S = \frac{24 \times 60}{\pi} \times \Phi_S \times RDR_{SE} [HA_S \sin \theta \sin \gamma + \cos \theta \cos \gamma HA_S]. \quad (10)$$

The reciprocal  $RDR_{SE}$  of the relative distance between the Sun and the Earth can be calculated by

$$RDR_{SE} = 1 + \frac{1}{30} \cos \left( \frac{2\pi}{365} \times DO \right). \quad (11)$$

The solar hour angle  $HA_S$  in formula (10) can be calculated by

$$HA_S = ar \cos [-\tan \theta \tan \gamma]. \quad (12)$$

The solar magnetic declination  $\gamma$  in formula (10) can be calculated by

$$\gamma = 0.4 \sin \left( \frac{2\pi}{365} DO - 1.4 \right). \quad (13)$$

Let  $MT_{\max-j}$  and  $MT_{\min-j}$  be the maximum and minimum absolute temperatures of month  $j$ . Then, the net long-wave radiation  $FS_C$  can be calculated by

The RBF centers were solved through clustering. Let  $B: \mathbb{R}^K \rightarrow \mathbb{R}^O$  be the mapping response of RBF;  $a_i$  the data of index sample  $i$  imported to the input layer, that is, the real-time data series of water demand indices for agricultural irrigation;  $A = (a_1, a_2, \dots)^T \in \mathbb{R}^K$  the input vector;  $b_k$  the output of output layer node  $k$ ;  $\omega_k$  the connection weight between all hidden layer nodes and output layer node  $k$ ;  $W_G$  the center  $G$  of hidden layer nodes;  $J$  and  $K$  the number of hidden layer centers and the total number of input layer nodes;  $F(\bullet)$  the basis function; and  $\|\cdot\|$  the Euclidean norm. Then, the RBFNN output can be expressed as

$$b_k = \sum_{i=1}^K \sum_{G=1}^J \omega_i(i, H) F \left( \|a_i - W_G\|^2 \right). \quad (19)$$

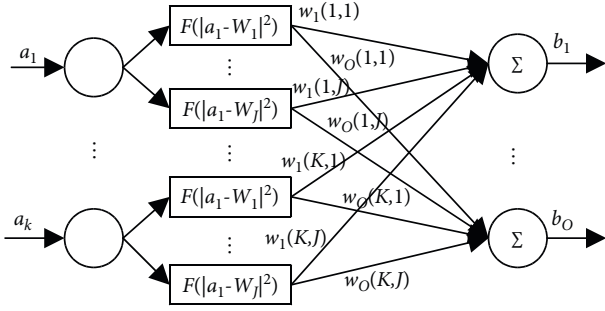


FIGURE 1: Structure of RBFNN.

The hidden layer output based on Gaussian function can be calculated by

$$c_i = \theta \left( \sum_{i=1}^K \sum_{G=1}^J \frac{(a_i - W_G)^2}{\gamma_G^2} \right). \quad (20)$$

To sum up, RBFNN has three adjustable parameters: variance, center position, and output layer weight. Under the premise that training index samples can reflect the essence of crop water demand prediction in the irrigated area,  $n_w$  uniformly distributed centers were selected with an interval of  $q$ , that is, network centers. On this basis, the variance  $s = q / (2n_w)^{-2}$  of the Gaussian function can be solved.

**3.2. Determination of Network Parameters.** The RBF centers were determined through fuzzy  $k$  means (FKM) clustering of unsupervised self-organizational learning. Let  $A_{US} = (A_1, A_2, \dots, A_N)$  be the set of  $N$  index samples imported for unsupervised learning;  $U$  the number of clusters;  $D = [d_{il}]_{U \times N}$  the membership matrix of sample  $l$  in cluster  $i$ ;  $d_{il}$  the membership function of sample  $l$  in cluster  $i$ ;  $O = [o_1, o_2, \dots, o_U]$  the set of cluster centers;  $\xi_{il}$  the distance from  $a_l$  to  $o_i$ ;  $\varepsilon$  the weighted index of the network; and  $\sigma$  the allowable error. To minimize the weighted sum of squares of the distance between sample center and cluster centers, an objective function of classification  $J(D, O) = \sum_{l=1}^N \sum_{i=1}^U (d_{il})^N (\xi_{il}^2)$  can be constructed and minimized. The detailed operations go as follows:

Step 1: initialize the network, set the initial number of iterations  $t = 0$ , and determine the initial values of  $N$ ,  $U$ , and  $D^{(t)}$ .

Step 2: derive the following from  $D^{(t)}$ :

$$o_i = \frac{1}{\sum_{l=1}^N (d_{il})^\varepsilon} \sum_{l=1}^N (d_{il})^N a_{il}, \quad i = 1, 2, \dots, U. \quad (21)$$

Step 3: suppose  $l = 1, \dots, N$ . The set of clusters satisfying the condition that the distance between sample center and cluster center is zero can be expressed as

$$Q_l = \{i | 0 \leq i \leq U, \xi_{il} = \|a_l - o_i\| = 0\}. \quad (22)$$

The set of the other clusters can be described by

$$Q^* = \{1, 2, \dots, U\} - Q_l. \quad (23)$$

If  $Q_l$  is an empty set, then

$$d_{il} = \frac{1}{\sum_{j=1}^U (\xi_{il} / \xi_{jl})^{2/(\varepsilon-1)}}. \quad (24)$$

If  $Q_l$  is not an empty set, then  $d_{il} = 0$ ,  $i$  belongs to set  $Q_l$ , and  $\sum_{i \in Q_l} d_{il} = 1$ . Make  $l = l + 1$ , and return to the start of Step 3.

Step 4: if  $\|D^{(t)} - D^{(t+1)}\|$  is smaller than the learning convergence rate  $\eta$ , then terminate the network training. If  $\|D^{(t)} - D^{(t+1)}\|$  is greater than  $\eta$ , add one to the number of iterations and return to the start of Step 2.

The hidden node width of RBFNN was determined by the  $k$ -nearest neighbors (KNN) algorithm. First, the sum of the distances from a sample center to the centers of the nearest  $K$  clusters was averaged. Then, the width was derived from the mean value. Let  $o'_j$  be the sample center closest to cluster center  $o_i$ . Then, the width of hidden layer nodes can be expressed as

$$HW_i = \frac{1}{l} \sum_{j=1}^l \|o_i - o'_j\|. \quad (25)$$

Let  $\text{Gauss}(a)$  be the Gaussian function. Then, the RBF can be approximated by the following function:

$$AF(a) = \sum_{i=1}^U \omega_i \times \text{Gauss}(\|a - o_i\|). \quad (26)$$

Solving the approximation function is equivalent to minimizing the following cost function by adjusting the weight set  $\{\omega_i | i = 1, 2, \dots, U\}$ :

$$\begin{aligned} \Phi(AF^*) &= \sum_{l=1}^N [\xi_l - AF(a_l)]^2 + \mu \|NOR^*\|^2 \\ &= \sum_{l=1}^N \left[ \xi_l - \sum_{i=1}^g \omega_i \text{Gauss}(\|a_l - o_i\|) \right]^2 + \mu \|NOR^*\|^2, \end{aligned} \quad (27)$$

where the first term is the standard error and the second term is the regularization term to satisfy the requirements on continuity and smoothness. Formula (27) can be sorted out as

$$\begin{aligned} \Phi(U^*) &= \sum_{l=1}^N [\xi_l - \text{Gauss} \cdot \omega]^2 + \mu (NOR^*, NOR^*)_M^2 \\ &= \|\xi_l - \text{Gauss} \cdot \omega\|^2 + \mu \omega^T H_0 \omega, \end{aligned} \quad (28)$$

where

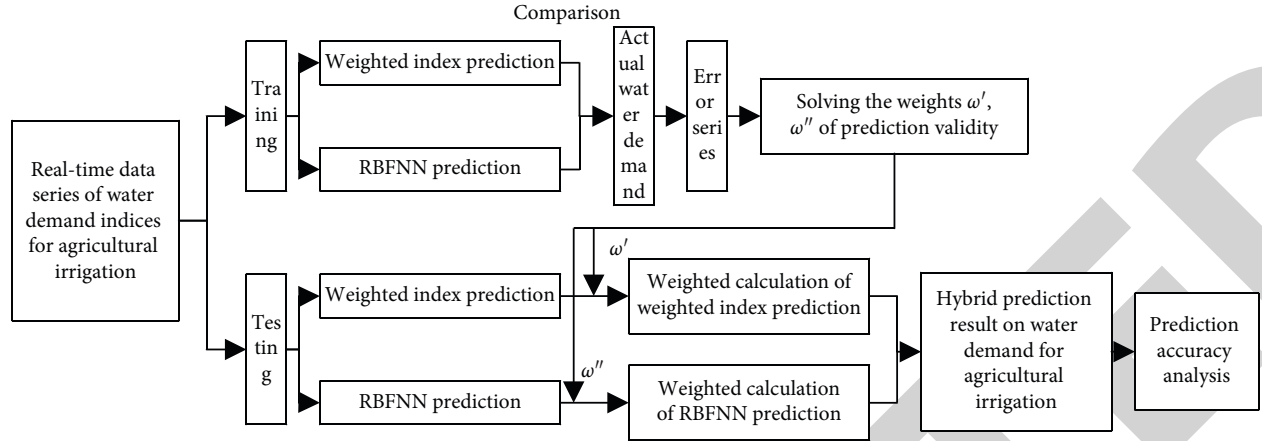


FIGURE 2: Steps of our hybrid prediction model.

$$\begin{aligned} \xi &= (\xi_1, \xi_2, \dots, \xi_N)^T; \\ \omega &= (\omega_1, \omega_2, \dots, \omega_U)^T, \\ \text{Gauss} &= \begin{bmatrix} g(a_1, o_1) & g(a_1, o_2) & \dots & g(a_1, o_U) \\ g(a_2, o_1) & g(a_2, o_2) & \dots & g(a_2, o_U) \\ \vdots & \vdots & \dots & \vdots \\ g(a_N, o_1) & g(a_N, o_2) & \dots & g(a_N, o_U) \end{bmatrix}. \end{aligned} \quad (29)$$

Minimizing  $\Phi(AF^*)$ ,

$$(g^T g + \mu g_0) \omega = g^T \xi. \quad (30)$$

The solution of the RBF approximation function should be normalized by

$$\omega = (g^T g + \mu g_0)^{-1} g^T \xi. \quad (31)$$

After setting  $l=l+1$ , the network iteration can be described by

$$\omega_{l+1} = \omega_l - \chi \frac{\gamma \Phi(U)}{\gamma \omega_k} = \omega_l - \chi [(g^T g + \mu g_0) \omega_l - g^T \xi]. \quad (32)$$

**3.3. Construction of Hybrid Prediction Model.** For accurate prediction of the water demand of agricultural irrigation, this paper combines RBFNN with an autoregression model into a water demand prediction model. In the model, the coefficient of the autoregression function can be approximated by the RBFNN.

Let  $u_b$  be the order of the model;  $O_i^j$  the center of RBFNN index samples;  $\mu_l$  the scale ratio;  $v_0$  and  $v_{bi}$  the state coefficients of the model;  $s_i^0$  and  $s_{il}^s$  the vector constants of the model;  $A(t-1)$  the index sample determining the output state of the time-varying model; and  $NO(t)$  the white noise. Then, the proposed prediction model can be established as

$$\begin{cases} b(t) = v_0 \left( A(t-1) + \sum_{i=1}^{u_b} v_{bi} (A(t-1)) b(t-i) + NO(t) \right), \\ v_0 (A(\tau-1)) = s_0^0 + \sum_{l=1}^N e^{-\mu_l^b} \|A(\tau-1) - O_l^b\|_2^2, \\ v_{bi} (A(\tau-1)) = s_{i0}^b + \sum_{l=1}^N s_{il}^b e^{-\mu_l^b} \|A(\tau-1) - O_l^b\|_2^2, \\ O_l = (O_{l,1}, O_{l,2}, \dots, O_{l,N})^T. \end{cases} \quad (33)$$

The weighted prediction result obtained from the water demand indices in Section 2 was combined with the prediction result of the RBFNN in Section 3, using validity weights. Figure 2 explains the steps of our hybrid prediction model. Firstly, the real-time data series of water demand indices for agricultural irrigation was divided into a training set and a test set. The two sets were used separately to predict

the water demand. Then, the predicted water demand was compared with the actual water demand of the agricultural irrigated area. The error between them was reduced to realize intelligent control of agricultural irrigation. Then, the validity weights  $\omega'$  and  $\omega''$  of the two methods were calculated.

The core of the intelligent control system for agricultural irrigation is a single-chip microcomputer (SCM) connected

to a triode amplifier circuit, which drives a relay. The opening/closing of the solenoid valve is controlled by the contact of the relay. Because of the small power of the selected solenoid valve, the triode type being adopted can stabilize the voltage and current and ensure the reliable functioning of the system. The soil humidity sensor transmits the humidity to the irrigation monitoring control system. The read humidity is compared with the preset lower bound of humidity to control the level at the pins of the SCM, in order to execute the irrigation mode.

During the intelligent control of actual agricultural irrigation,  $\omega'$  and  $\omega''$  change with the updates of the control signals of the irrigation solenoid valve, in order to improve the precision of irrigation control.

Similarly, the test set was predicted separately by two methods. Then, the final hybrid prediction result was obtained through the weighted calculation, using the validity weights  $\omega'$  and  $\omega''$ . The final step is to evaluate the accuracy of water demand prediction.

#### 4. Intelligent Control Strategy Based on Water Demand Prediction

Based on the hybrid prediction of water demand for agricultural irrigation, the authors developed a precision irrigation intelligent control system (Figure 3), which encompasses a distributed wireless sensor subsystem, a water demand prediction subsystem, and an agricultural irrigation control subsystem. The agricultural irrigation control subsystem evaluates whether the target crop needs water by comparing soil humidity with predicted crop water demand. If the crop needs water, the solenoid valve will be launched for irrigation. The duration of irrigation will be controlled by the water quota for agricultural irrigation. Figure 4 shows the signal conversion in the distributed wireless sensor subsystem, which measures soil humidity in real time, and the agricultural irrigation control subsystem, which predicts the indices of crop water demand.

The analog signal from the wireless soil humidity sensor has a low voltage. After passing through the amplifier circuit, the voltage is increased to the useable level. Then, the analog signal is converted by the converter into a digital signal. After that, the SCM level is transformed by the chip into the level for the serial port of the host, which supports intelligent control.

The precision irrigation intelligent control system should operate rapidly while occupying a small memory. For this purpose, Keil  $\mu$  Vision 4.0, which contains efficient C compilers and project managers, was selected to compile the program code of the system. The main program flow is presented in Figure 5. Once the intelligent control system starts, the main program will be initialized, and the command will be read. When the water demand surpasses the actual soil humidity, the solenoid valve will be opened to start sparkling irrigation in the irrigated area. The irrigation will last until the end of the preset duration. Then, the valve will be closed to terminate the irrigation.

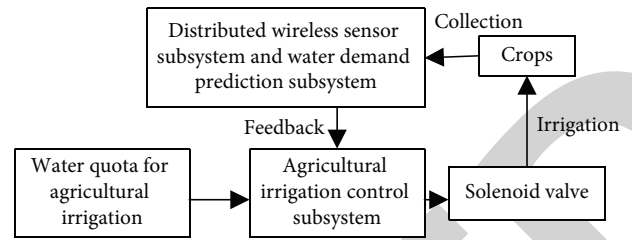


FIGURE 3: Structure of intelligent control system.

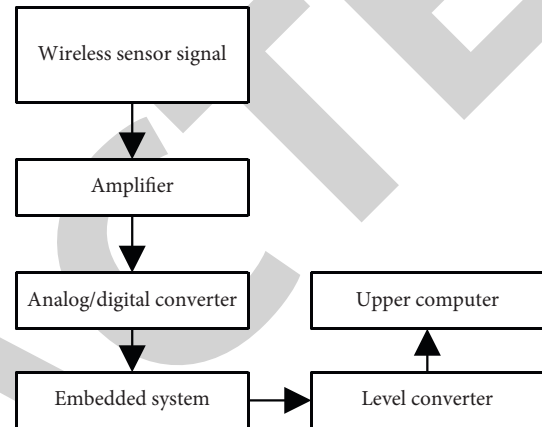


FIGURE 4: Signal conversion in subsystems.

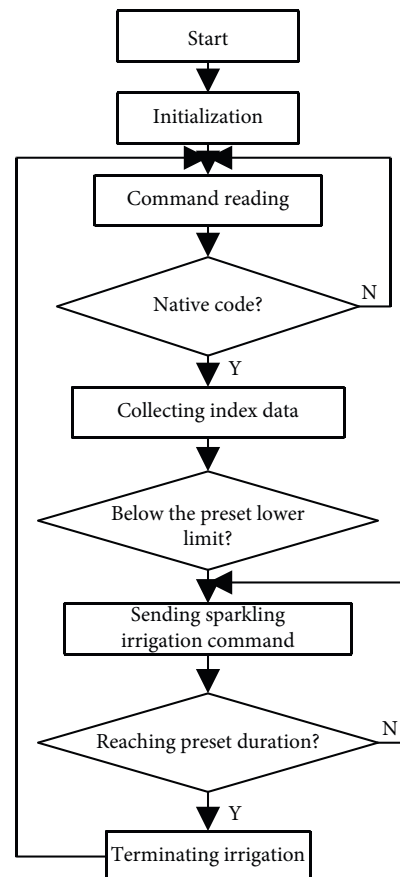


FIGURE 5: Main program flow.

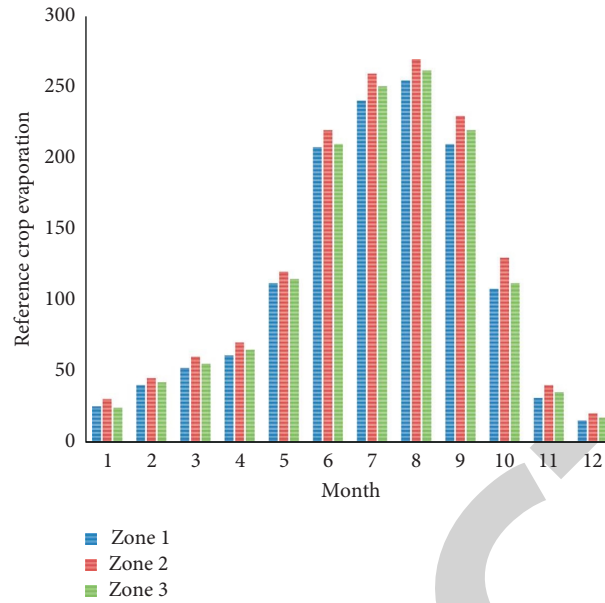


FIGURE 6: Monthly mean evaporations of different zones in an irrigated area.

TABLE 1: Ecological water demand obtained by area quota method.

Crop	Irrigation quota	Current state			Predicted state		
		Area	Calculated water demand	Actual water demand	Area	Calculated water demand	Actual water demand
Soybean	5,100	35,712	1.856	1.235	78,512	4.371	3.464
Corn	4,960	89,643	4.182	2.372	93,256	4.325	3.375

## 5. Experiments and Results Analysis

In the agricultural irrigated area, the phreatic evaporation of crops is greatly affected by the actual atmospheric evaporation capacity. Figure 6 presents the monthly mean evaporations of different planting zones in the agricultural irrigated area, which were derived from the daily observation data. It can be inferred that the crop evaporation was relatively large from June to September. The change trend meets the actual variation of water demand.

During the estimation of the ecological water demand of vegetation, the key of the area quota method lies in determining the ecological water demand quota of each type of plant. In the target irrigated area, the irrigation quotas for soybean and corn are 5,100 and 4,960  $\text{m}^3/\text{hm}$ , respectively. Considering the planting areas and expansion goals of the crops in the current year, the area quota method was employed to compute the ecological water demands of soybean and corn in the irrigated area. The results (Table 1) show that the current ecological water demand of soybean and corn in the target irrigated area was 1.856 and 4.182  $\text{m}^3$ , respectively. The total water demand reached 6.028  $\text{m}^3$ . Under the expansion goals, the ecological water demand of soybean and corn was predicted as 4.371 and 4.352  $\text{m}^3$ , respectively. The total water demand reached 8.723  $\text{m}^3$ . The traditional agricultural irrigation control systems are very inefficient in water use. The irrigation water utilization coefficient is as low as 0.6. Based on this coefficient, the

actual ecological water demand of soybean and corn in the target irrigated area was 1.235 and 2.372  $\text{m}^3$ , respectively. The total water demand reached 3.607  $\text{m}^3$ . By contrast, our precision irrigation intelligent control system can elevate the utilization coefficient to 0.85. On this basis, the actual ecological water demand of soybean and corn in the target irrigated area was 3.464 and 3.375  $\text{m}^3$ , respectively. The total water demand reached 6.839  $\text{m}^3$ .

By the phreatic evaporation method, the ecological water demand of crops per unit area in the target irrigated area averaged at 2,749  $\text{m}^3/\text{hm}$ . The predicted values under the expansion goals are given in Table 2. It can be seen that the total annual ecological water demand of soybean in the irrigated area was 2.322  $\text{m}^3$ , and that of corn was 2.573  $\text{m}^3$ . The total water demand reached 4.895  $\text{m}^3$ .

By the vegetation evapotranspiration method, the mean ecological water demand of crops per unit area in the target irrigated area was 3,359  $\text{m}^3/\text{hm}$ . The predicted values under the expansion goals are given in Table 3. It can be seen that the total annual ecological water demand of soybean in the irrigated area was 2.841  $\text{m}^3$ , and that of corn was 2.597  $\text{m}^3$ . The total water demand reached 5.438  $\text{m}^3$ .

Figures 7 and 8 compare the predicted water demands for the target irrigated area and prediction errors, respectively. Table 4 sums up the absolute errors of water demand prediction by different methods. The error metrics include maximum absolute error, mean absolute error, and root mean square of absolute error. From Figures 7 and 8 and Table 4, it



TABLE 2: Ecological water demand obtained by the phreatic evaporation method.

Crop	Ecological water demand						
	January	February	March	April	May	June	Full year
Soybean	0.023	0.051	0.139	0.278	0.356	0.368	2.322
Corn	0.027	0.056	0.161	0.306	0.375	0.409	2.573
	July	August	September	October	November	December	
Soybean	0.351	0.306	0.224	0.152	0.051	0.023	
Corn	0.401	0.358	0.246	0.143	0.069	0.022	

TABLE 3: Ecological water demand obtained by vegetation evapotranspiration method.

Crop	Ecological water demand						
	January	February	March	April	May	June	Full year
Soybean	0.012	0.023	0.065	0.113	0.518	0.532	2.814
Corn	0.021	0.053	0.168	0.305	0.358	0.409	2.597
	July	August	September	October	November	December	
Soybean	0.560	0.473	0.251	0.182	0.062	0.023	
Corn	0.401	0.358	0.260	0.153	0.076	0.035	

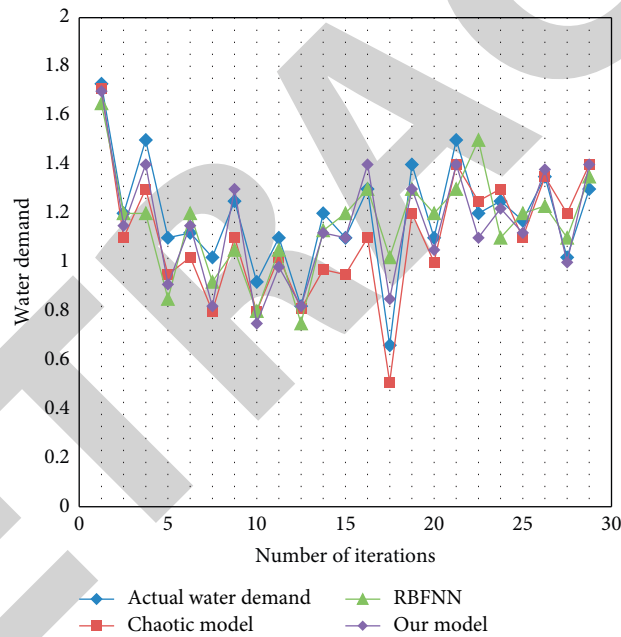


FIGURE 7: Comparison of predicted water demands.

TABLE 4: Absolute errors of water demand prediction.

	Method	Maximum	Mean	Root mean square
Water demand prediction	Chaotic method	0.3	0.115	0.1235
	RBFNN	0.15	0.0965	0.1053
	Our hybrid prediction method	0.13	0.028	0.0342

can be inferred that the chaotic model, traditional RBF model, and our hybrid prediction model all achieved high accuracy in forecasting the water demand in the irrigated area. However, the three error metrics of our model were way smaller than

those of other models. Therefore, our method can adapt well to the index value fluctuations in prediction and realize stable prediction performance, compared with the chaotic model and traditional RBF model.



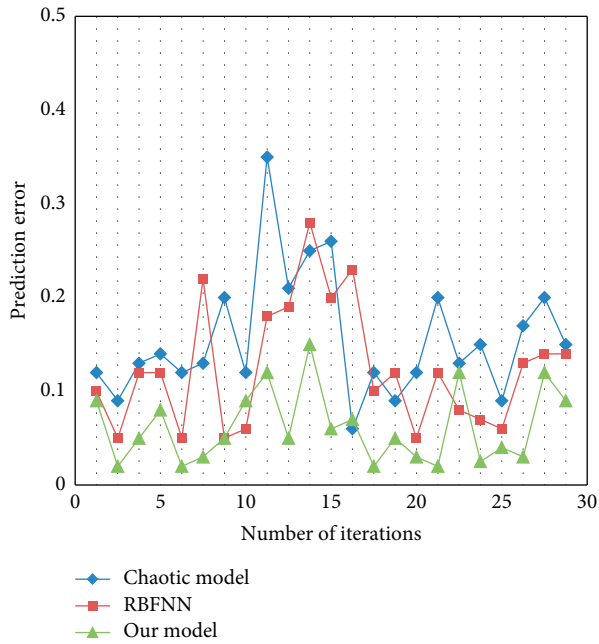


FIGURE 8: Comparison of prediction errors.

## 6. Conclusions

This paper presents an intelligent control method for agricultural irrigation through ANN-based water demand prediction. Firstly, the ecological water demand for agricultural irrigation of crops was calculated, and an RBFNN was constructed for predicting the water demand of agricultural irrigation. Then, the authors developed an intelligent control strategy for agricultural irrigation based on water demand prediction, detailed the structure of the intelligent control system, and designed the main program. Through experiments, the ecological water demand for crops in the targeted irrigated area was separately calculated by the area quota method, the phreatic evaporation method, and the vegetation evapotranspiration method, providing the data support to the weighted prediction of water demand indices. Then, the authors compared the predicted water demands for the target irrigated area and prediction errors, respectively, and demonstrated that our method could predict the water demand in the irrigated area more accurately and stably than other models.

This paper preliminarily explores and discusses the calculation of agricultural irrigation water demand. There are several limitations of the work, which need to be improved in the future. (1) Due to data availability, the meteorological data are the mean of the three stations around the study area. Thus, the calculated evaporation may deviate from the actual evaporation in the study area. (2) The correlation between groundwater level and vegetation area is not fully demonstrated. The vegetation response to groundwater level needs to be deeply analyzed in the future.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This paper was supported by the Open Research Fund Program of State Key Laboratory of Hydrosience and Engineering of Tsinghua University (Grant no. sklhse-2020-A-01).

## References

- [1] C. Ramirez, Y. Almulla, and F. F. Nerini, "Reusing wastewater for agricultural irrigation: a water-energy-food Nexus assessment in the North Western Sahara Aquifer System," *Environmental Research Letters*, vol. 16, no. 4, Article ID 044052, 2021.
- [2] S. Khaeez and S. M. H. Shahdany, "Non-structural modification of agricultural water distribution systems in large scale irrigation districts," *Computers and Electronics in Agriculture*, vol. 184, Article ID 106102, 2021.
- [3] A. B. Ashu and S.-I. Lee, "The effects of climate change on the reuse of agricultural drainage water in irrigation," *KSCE Journal of Civil Engineering*, vol. 25, no. 3, pp. 1116–1129, 2021.
- [4] S. Mitra, S. Singh, and P. Srivastava, "Sensitivity of groundwater components to irrigation withdrawals during droughts on agricultural-intensive karst aquifer in the Apalachicola-Chattahoochee-Flint River Basin," *Journal of Hydrologic Engineering*, vol. 24, no. 3, Article ID 05018032, 2019.
- [5] S. K. Mandal, S. K. Dutta, S. Pramanik, and R. K. Kole, "Assessment of river water quality for agricultural irrigation," *International Journal of Environmental Science and Technology*, vol. 16, no. 1, pp. 451–462, 2019.
- [6] A. Selmani, H. Oubehar, M. Outanoute et al., "Agricultural cyber-physical system enabled for remote management of solar-powered precision irrigation," *Biosystems Engineering*, vol. 177, pp. 18–30, 2019.
- [7] I. P. Kruzhilin, T. A. Gamm, A. A. Mushinskiy, and S. D. Fomin, "Assessing the ecological state of agricultural irrigated fields of the Orenburg gas processing complex with dumping sewage water for crop irrigation," *IOP Conference Series: Earth and Environmental Science*, vol. 350, no. 1, Article ID 012037, 2019.
- [8] A. Christou, P. Karaolia, E. Hapeshi, C. Michael, and D. Fatta-Kassinou, "Long-term wastewater irrigation of vegetables in real agricultural systems: concentration of pharmaceuticals in soil, uptake and bioaccumulation in tomato fruits and human health risk assessment," *Water Research*, vol. 109, pp. 24–34, 2017.
- [9] H. Sun, S. Wang, and X. Hao, "An improved analytic hierarchy process method for the evaluation of agricultural water management in irrigation districts of North China," *Agricultural Water Management*, vol. 179, pp. 324–337, 2017.
- [10] H. I. Essaid and R. R. Caldwell, "Evaluating the impact of irrigation on surface water - groundwater interaction and stream temperature in an agricultural watershed," *The Science of the Total Environment*, vol. 599–600, pp. 581–596, 2017.
- [11] C. Sutcliffe, J. Knox, and T. Hess, "Managing irrigation under pressure: how supply chain demands and environmental objectives drive imbalance in agricultural resilience to water shortages," *Agricultural Water Management*, vol. 243, Article ID 106484, 2021.

## Retraction

# Retracted: A Data-Driven Control Strategy for Urban Express Ramp

### Computational Intelligence and Neuroscience

Received 31 October 2023; Accepted 31 October 2023; Published 1 November 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] L. Zhang, M. Zhang, J. Ma, and J. Ge, "A Data-Driven Control Strategy for Urban Express Ramp," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 6540972, 12 pages, 2021.

## Research Article

# A Data-Driven Control Strategy for Urban Express Ramp

Liyan Zhang , Min Zhang , Jian Ma , and Jing Ge 

School of Civil Engineering, Suzhou University of Science and Technology, 1701 Binhe Road, New District, Suzhou, Jiangsu 215011, China

Correspondence should be addressed to Jian Ma; 9764634@qq.com

Received 13 October 2021; Accepted 2 November 2021; Published 16 November 2021

Academic Editor: Daqing Gong

Copyright © 2021 Liyan Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Expressway, as the main artery of urban traffic, realizes the smooth operation of the whole urban road network through reasonably balancing the traffic flow. However, due to the lack of reasonable and effective traffic control, the safety and congestion of expressways are becoming more and more serious. The development of intelligent network technology provides a new idea to solve the control problem of expressways. In this paper, a data-driven ramp control model of urban expressway is constructed. The interaction of traffic information is realized through intelligent network connection technology. The cooperative control strategy of VSL and RM is adopted. The mutual feedback of VSL and RM is realized based on the improved METANET model. The simulation experiment based on VISSIM secondary development shows that the collaborative control strategy under the intelligent network environment could make the vehicle travel time reduced by 20.59% and the speed difference between adjacent sections of the expressway mainline by 34.07%, which realized the coordinated control of the mainline and the on-ramp under the intelligent network environment, alleviate the expressway traffic congestion, reduce the traffic pressure, and improve the efficiency of the road network.

## 1. Introduction

The sustained economic development has promoted the rapid growth of the transportation industry. The surge of motor vehicle ownership and urban traffic demand has also derived various traffic problems. Congestion has become a major problem in urban governance. Among them, the transportation efficiency of expressway has a significant impact on the smooth passage of urban traffic network. Due to the influence of urban traffic and road network, expressways have the characteristics of dense entrance and exit ramps, complex road conditions, and low speed limits [1]. How to improve the operation efficiency of expressways through the implementation of reasonable traffic control has become the focus of scholars at home and abroad.

The research scope of expressway abroad is very wide, including planning, design, operation, management, and other aspects. Ramp control method first appeared in the United States in 1963 and achieved good control effect at that time, which can improve the operation speed of road traffic flow and alleviate traffic congestion [2]. Since then, a

large number of scholars have carried out ramp metering (RM) and variable speed limit control (VSL). The scope of RM includes on-ramp and off-ramp. This study focuses on the analysis of on-ramp control methods. Carlson designed a simple local mainline traffic flow feedback controller; that is, the mainline traffic flow control (MTFC) is realized through VSL. This method does not need traffic demand prediction and is convenient for on-site implementation [3]. Walraven et al. used reinforcement learning to control the maximum allowable speed on the expressway, so as to optimize the mainline traffic flow. A series of simulation experiments show that this strategy can significantly reduce the congestion under high traffic demand, and the introduction of traffic prediction can also improve the control effect. At the same time, this strategy has sufficient robustness for speed and density control [4]. The research proposed above is only for a single ramp, ignoring the overall control of the road network. Carlson et al. combined RM and MTFC through VSL, evaluated its effect by simulation software, and gave the optimal control method by comparing with different control methods [5]. Han et al. proposed a model-based network

coordinated control method. The upper layer of the control system optimizes the total travel time of the system by controlling the total flow of the on-ramp into the mainline, and the lower layer distributes the optimal total flow to each on-ramp of the expressway. The robustness of the system is verified by simulation. The results show that it has less calculation and is suitable for field implementation [6].

Domestic scholars have also made some achievements in the integrated control of urban roads and expressways. Yu et al. proposed an optimal control model for the conditions of implementing VSL in expressway congestion areas and used simulation to prove that, compared with fixed speed limit, this method can reduce the system travel time by 7.45% [7]. Lian introduced CMEM model from the perspective of traffic benefit and environmental benefit and compared the application effect of joint control model through simulation [8]. The difficulty of traffic network integrated control is that when multiple control strategies are adopted at the same time, their mutual effects should be considered; otherwise the control effect will be affected and the traffic network will deteriorate.

To sum up, foreign scholars began to explore ramp control earlier and have many relatively mature control algorithms. Although domestic research started late, it has also made some achievements. Compared with the research on single ramp control, the research on multiramp coordinated control and road network linkage control is less, and the ramp control system of urban expressway still needs further research. At present, the ramp control strategy of expressway is lack of research on real-time control of vehicle flow under different traffic conditions.

The rise and development of intelligent networking technology provides a new idea for solving the problem of ramp control of urban expressway. Intelligent networked transportation system [9] refers to a joint control system formed by cloud computing, big data, and Internet of things technologies. With the continuous advancement of artificial intelligence and network communication technology research, automobile intelligence and network connection have become global research focus. The definition of intelligent connected car is equipped with advanced on-board sensors, controllers, actuators, and other devices and integrates modern communication and network technology to realize vehicle to everything (V2X) intelligent information exchange and sharing, complex environment perception, intelligent decision-making, and collaboration of a new generation of cars with functions such as control, safe, and efficient driving, and ultimately replacing human operation [10, 11]. In the intelligent networking environment, wireless communication technology can be used to complete information exchange between vehicles, vehicles and roadside facilities, which helps managers obtain accurate road conditions in real time and realize accurate control of vehicles.

At present, the research on ramp control in intelligent network environment is still in the exploratory stage. The theory and practical application are far from mature as the classical ramp control strategy. Foreign researches on ramp control in intelligent network environment mostly focus on the control of vehicle merging, the safety

influencing factors of vehicle merging, and related simulation evaluation. Sakaguchi et al. proposed the import control algorithm based on virtual vehicle and the data transmission method of vehicle to vehicle. Lu et al. proposed a vehicle safety import control algorithm based on wireless communication. Qiao et al. used a deep Q network to build a reinforcement learning model to solve the problem of unmanned car ramp merging. These methods mainly discussed the single-point ramp merging control from the perspective of workshop communication and bicycle control and lacked control of vehicles in the ramp area, coordination, and guidance of traffic flow [12]. Zhang proposed a method to control the on-ramp traffic flow of an expressway under the condition of vehicle-road coordination and conducted simulation experiments on the mainline and the ramp in different traffic conditions. The results show that the method can reduce the on-ramp delay in the peak state by 19.2%, and traffic conflicts are reduced by 32.3% [13]. Luo and Jiang proposed an optimization algorithm for ramp merging trajectory in a networked environment. The simulation experiment proved that the proposed model can reduce the total delay of vehicles by 59.7% and realized that vehicles can pass the ramp traffic area at a higher speed [14]. Miao proposed a multiramp collaborative control framework in a vehicle-road collaborative environment and then established a simulation example based on the Paramics simulation platform and analysed the response sensitivity of control cycle, distance cost, and control rate, but the control strategy is mainly through traffic guidance in a larger range that is realized without precise control of bicycles [15]. At present, the research on ramp control and simulation under the environment of intelligent networking focuses on the microresearch on the merging control of ramp vehicles and lacks coordinated control of vehicles in the ramp area of expressways. The reason is that the existing mainstream traffic simulation software does not fully consider the impact of the intelligent network environment on the traffic state and traditional control methods in the ramp area. It is necessary to realize the intelligent network environment and related logic control through secondary development or independent development, which has high requirements for researchers' program design and writing. Therefore, it is of great significance to study the ramp control in the intelligent network environment.

Based on the above analysis, the focus of this paper is to study the cooperative control of vehicles in the ramp area of expressways with the help of intelligent network technology. The framework of expressway ramp control system and the expressway ramp collaborative control model in intelligent network environment is proposed. Considering the cooperative control of mainline VSL and RM, the dynamic change process of mainline traffic flow on expressway is described and predicted based on the improved METANET model. Through the secondary development of simulation software, the reliability of the proposed control model is verified, which lays a foundation for further research.



## 2. Expressway Ramp Control System in Intelligent Networked Environment

*2.1. Communication Requirements of Intelligent Networked Systems.* Traffic information is the core content of the intelligent networked system. The expressway ramp control system in the intelligent networked environment puts forward further requirements for the perception and interaction of traffic information.

*2.1.1. Traffic Information Perception.* In the traditional driving environment, the driver mainly obtains external traffic information through sensory organs, such as vehicle operating conditions and environmental information, road instruction information, etc. In the intelligent networked environment, information can also be obtained through on-board sensors, etc., and information can be shared between vehicles, which greatly reduces the error of traditional information perception. Traffic information perception is the basis of information interaction. It can provide data sources for intelligent networked systems. The specific information it collects includes perception of various operating conditions in vehicle operation; perception of expressway environment; vehicles and objects perception; perception of vehicle location, etc.

*2.1.2. Traffic Information Interaction.* Traffic information interaction is a process of information transmission, sharing, and exchange. Through information interaction, each node of the Internet of Vehicles can independently obtain information about the environment and other nodes [16]. The communication of the intelligent network system mainly includes vehicle-vehicle (V2V), vehicle-infrastructure (V2I), vehicle-pedestrian (V2P), and vehicle-external network (V2N) [17]. V2X is the key technology of the future intelligent transportation system. It enables real-time data exchange between vehicles and the environment, thereby obtaining real-time traffic data information, alleviating congestion, reducing accidents, and improving the overall efficiency of the road network.

*2.2. The Overall Framework of the Expressway Ramp Control System.* This paper proposes an urban expressway ramp control system in an intelligent networked environment. Its control structure is shown in Figure 1. It consists of the mainline, on-ramp, and vehicle control subsystems. The control goal of the expressway mainline control subsystem is to implement VSL on the upstream vehicles of the mainline to relieve the bottleneck of the expressway. The system consists of the mainline, the ramp side unit RSU, on-board equipment, and the traffic information management control centre. The control objective of the on-ramp control system is to restrict the number of vehicles entering the expressway during the control period and reduce traffic delays in the confluence area. The system consists of the mainline, on-ramp RSU, on-ramp queuing detector, on-ramp and traffic information management control centre, and so on. The on-

board control subsystem is composed of vehicles equipped with on-board equipment in the ramp control area. Its main function is to send vehicle information to the roadside unit while receiving instructions from the ramp layer control system and pass through the ramp in turn according to the given traffic strategy.

*2.3. On-Ramp Control System.* The on-ramp control system is composed of the mainline, on-ramp RSU, on-ramp queuing detector, and on-ramp and traffic information management control centre. The function of the mainline and on-ramp RSU is to obtain real-time traffic status information on the upstream section of the mainline and the on-ramp. At the same time, the on-board equipment responds to the demand to send vehicle driving information to the mainline and on-ramp RSU, and the RSU transmits the received information to the on-ramp control centre in real time. The on-ramp control centre determines the on-ramp control strategy by sharing information with the traffic information management control centre, combined with the mainline operation status, and sends this control command to the on-ramp vehicle through the on-ramp RSU. After receiving the instruction, the on-board equipment adjusts the driving behavior of the vehicle and adopts slowing down or stopping and waiting. The function of the on-ramp queuing detector is to detect the length of the ramp queuing and send the detection information to the on-ramp control centre in time. When the ramp queuing has reached the maximum queuing length and the expressway confluence area is blocked, the ramp control centre needs to issue instructions in time to stop the flow of traffic and close the entrance ramp (see Figure 2).

## 3. Cooperative Control Model in the Intelligent Networked Environment

*3.1. The Collaborative Control Process of Expressway Ramps in the Environment of the Internet of Vehicles.* The expressway ramp cooperative control process is specifically described as follows:

- (1) Use the traffic flow model to predict the traffic state of the mainline of the expressway.
- (2) Set the objective function and constraint conditions of the cooperative control system and obtain the optimal ramp regulation rate and variable limit rate through NSGA-II Algorithm solving and act on the system.
- (3) According to the obtained optimal ramp adjustment rate, adjust the ramp inflow to perform ramp control.
- (4) Adjust the vehicle's desired speed according to the obtained optimal variable speed limit for mainline speed guidance control.
- (5) Run the simulation to obtain the traffic flow state after the collaborative control. In the next rolling cycle, repeat the above steps to achieve rolling optimization.

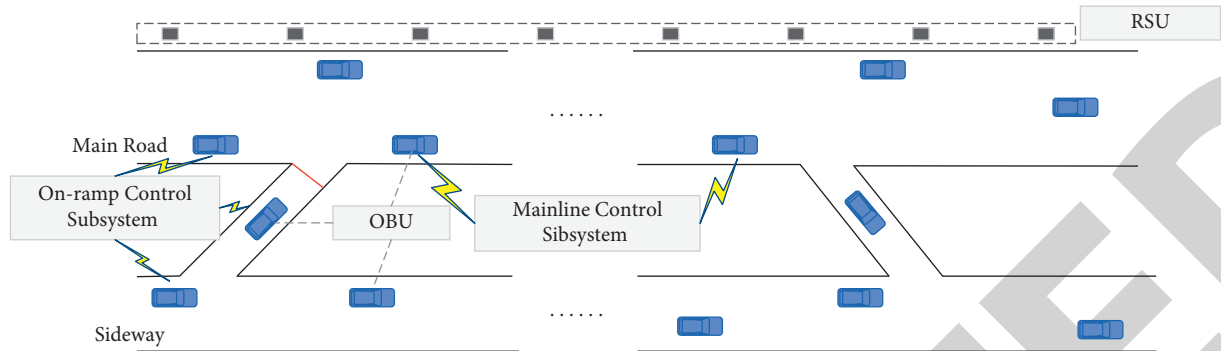


FIGURE 1: Schematic diagram of the expressway ramp control system in the intelligent network environment.

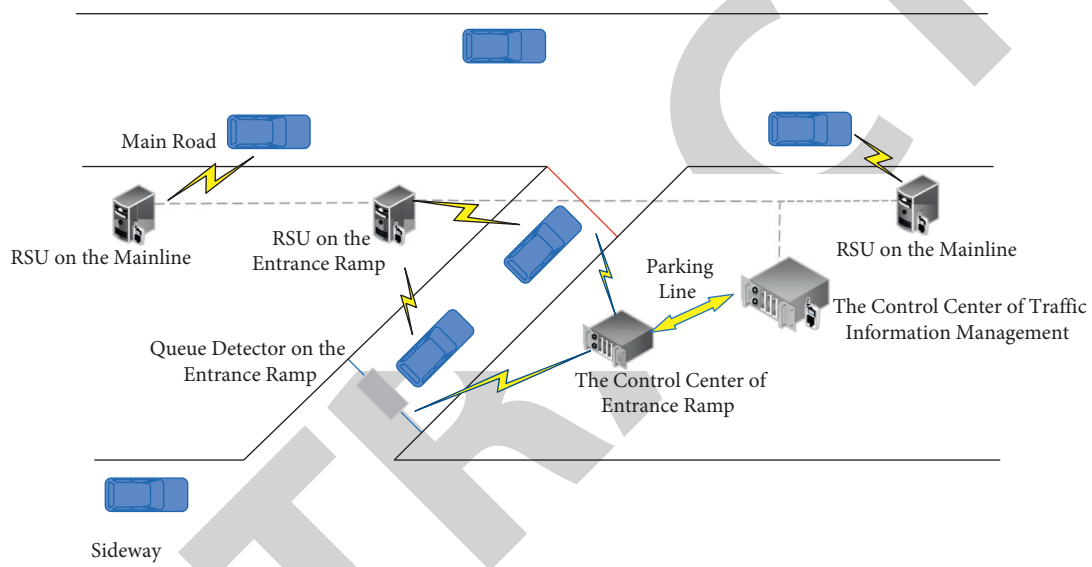


FIGURE 2: Schematic diagram of the on-ramp control system.

### 3.2. Improved METANET Model

**3.2.1. METANET Basic Model.** According to the accuracy of the description of the road state by the traffic flow model, it is divided into three types: macromodel, mesomodel, and micromodel. The micromodel is to study the law of traffic flow at a certain point or section in a short period of time. Currently, the commonly used micromodels include car following and lane changing models; the macromodel mainly discusses the relationship between the three elements of traffic flow (flow, speed, and density) and uses it as a basis to describe the general traffic characteristics of the road, which is specifically expressed as the temporal and spatial changes and distribution laws of the traffic flow state [18]; the mesoscopic model is between the two, and it has less calculations than the microscopic model and is more expensive than the macroscopic model. High accuracy [19] is suitable for analysing regional traffic conditions. After more than half a century of continuous research, the traffic flow model has developed to a relatively complete level, which can accurately describe the real behavior of traffic flow [20].

At present, the more typical macromodels describing the state of traffic flow include LWR model [21], CTM model [22], and METANET model [23]. Compared with the LWR model and the CTM model, the METANET model considers the dynamic change process of speed more comprehensively. This model is also the most widely used macroscopic model. Therefore, this study uses the METANET model to describe the dynamic process of traffic flow.

The METANET model can be used to describe the macroscopic traffic flow operating state of the road [24]. The main idea is to discretize the traffic wave, divide the expressway traffic flow into several nodes and road sections, and predict the traffic state of each road section at the next moment based on the current traffic state of each road section and the connection between adjacent road sections. In order to realize the estimation of traffic flow status and basic parameters [25], the model is divided into segments  $i$  as shown in Figure 3.

The METANET model mainly focuses on the relationship between flow, speed, and density. The basic model of the road section is as follows:



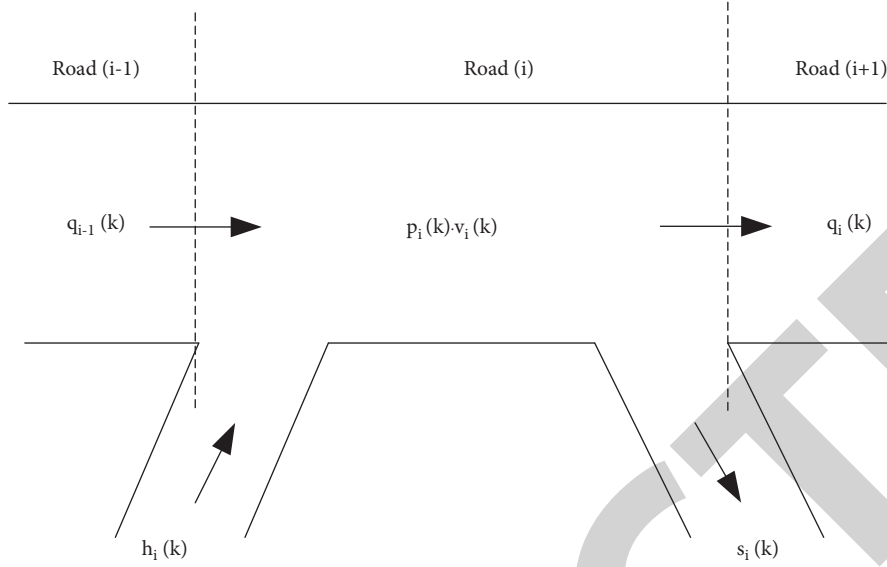


FIGURE 3: METANET traffic flow model unit diagram.

$$q_i(k) = \rho_i(k) \cdot v_i(k) \cdot \lambda_i, \quad (1)$$

$$\rho_i(k+1) = \rho_i(k) + \frac{T}{L_i \lambda_i} [q_{i-1}(k) - q_i(k) + h_i(k) - s_i(k)], \quad (2)$$

$$v_i(k+1) = v_i(k) + \frac{T}{\tau} \{H[\rho_i(k)] - v_i(k)\} + \frac{T}{L_i} v_i(k) [v_{i-1}(k) - v_i(k)] - \frac{f T}{\tau L_i} \frac{\rho_{i+1}(k) - \rho_i(k)}{\rho_i(k) + \kappa}, \quad (3)$$

$$H[\rho_i(k)] = v_f \cdot \exp\left[-\frac{1}{\alpha_i} \left(\frac{\rho_i(k)}{\rho_{cr}(k)}\right)^{\alpha_i}\right], \quad (4)$$

$$h_i(k) = \min\left\{\left[d_i(k) + \frac{\omega_i(k)}{T}\right], Q_i(k)\right\}, \quad (5)$$

$$\omega_i(k) = \omega_i(k-1) + T[d_i(k-1) - h_i(k-1)]. \quad (6)$$

In equations (1) to (6),  $\rho_i(k)$  is traffic density of road section  $i$  at time  $k$  (veh/km/lane);  $q_i(k)$  is the traffic flow of the road section  $i$  at the moment  $k$  (veh/h);  $h_i(k)$  is the inbound traffic volume of the section  $i$  at the moment  $k$  (veh/h);  $s_i(k)$  is the amount of traffic leaving the road section at the moment of (veh/h);  $T$  is the sampling period ( $h$ );  $L_i$  is length of road section  $i$  (km);  $\lambda_i$  is number of lanes in road section  $i$ ;  $v_i(k)$  is average traffic flow speed of section  $i$  at time  $k$  (km/h);  $H[\rho_i(k)]$  is desired speed (km/h);  $f$  is speed density sensitivity parameter;  $\tau$  is speed adjustment factor;  $\kappa$  is compensation coefficient to prevent  $\rho_i(k)$  from being too small;  $\rho_{cr}(k)$  is critical traffic flow density (veh/km/lane);  $\alpha_i$  is model coefficient;  $Q_i(k)$  is on-ramp capacity (veh/h);  $\omega_i(k)$  is the number of vehicles queuing at the entrance ramp at the moment of the road section  $i$ ;  $d_i(k-1)$  is the traffic demand of the on-ramp on the road section  $i$  at  $(k-1)$ , and 0 if there is no on-ramp.

The original METANET model is used to describe the state of the mainline traffic flow under collaborative control. Since the impact of the mainline VSL and RM on it is not discussed, the description is not accurate. Therefore, this article will comprehensively consider the effects of both to improve the original METANET traffic flow model.

### 3.2.2. METANET Model considering VSL

(1) *Desired Speed.* During the speed control of the mainline, it will have a certain impact on the speed of the expressway section. The expected speed needs to be corrected. The corrected expected speed model is as follows:

$$H[\rho_i(k)] = \min\left\{v_f \cdot \exp\left[-\frac{1}{\alpha_i} \left(\frac{\rho_i(k)}{\rho_{cr}(k)}\right)^{\alpha_i}\right], b_i(k)v_f\right\}, \quad (7)$$

where  $b_i(k)$  is variable speed limit, and value range is  $b_{\min,i}(k) \leq b_i(k) \leq 1$ ,  $b_{\min,i}(k) \in [0, 1]$ .

(2) *Velocity Density Sensitivity Coefficient.* The speed density sensitivity parameter  $f$  essentially means the partial derivative of the expected speed to the density. When the expected speed changes under the variable speed limit control, the speed density sensitivity parameter will also be affected. Therefore, the calculation of  $f_i(k)$  in this paper is as follows:

$$f_i(k) = -\frac{v_f}{\rho_{cr}(k)} \cdot \exp\left[-\frac{1}{\alpha_i} \left(\frac{\rho_i(k)}{\rho_{cr}(k)}\right)^{\alpha_i}\right]. \quad (8)$$

It can be seen from the above formula that the velocity density sensitivity coefficient  $f_i(k) < 0$ ; that is, the speed is inversely proportional to the density, which is in line with the basic theory of traffic flow.

(3) *Ramp Inflow.* The amount of ramp inflow is affected by many factors. Let  $\rho_{jam}$  be the congestion density; then the

amount of ramp inflow considering the influence of the mainline VSL of the expressway is  $Q_i(k)\rho_{jam} - \rho_i(k)/\rho_{jam} - \rho_{cr}$ . At this time, the ramp inflow model can be expressed:

$$h_i(k) = \min \left\{ \left( d_i(k) + \frac{\omega_i(k)}{T} \right), Q_i(k) \frac{\rho_{jam} - \rho_i(k)}{\rho_{jam} - \rho_{cr}} \right\}. \quad (9)$$

### 3.2.3. METANET Model considering RM

(1) *Mainline speed reduction.* When the on-ramp traffic flows into the mainline, it will reduce the speed of the mainline; set  $v_{r,i}(k)$  is the speed at which the ramp volume converges into the mainline. The vehicle speed reduction at the on-ramp is calculated as follows:

$$\Delta v_{r,i}(k) = \frac{\delta T v_i(k) h_i(k)}{L_i \lambda_i [\rho_i(k) + \kappa]}, \quad (10)$$

$$q_i(k) = \rho_i(k) \cdot v_i(k) \cdot \lambda_i, \quad (12)$$

$$\rho_i(k+1) = \rho_i(k) + \frac{T}{L_i \lambda_i} [q_{i-1}(k) - q_i(k) + h_i(k) - s_i(k)], \quad (13)$$

$$v_i(k+1) = v_i(k) + \frac{T}{\tau} \{ H[\rho_i(k)] - v_i(k) \} + \frac{T}{L_i} v_i(k) [v_{i-1}(k) - v_i(k)] - \frac{f_i(k)}{\tau} \frac{T}{L_i} \frac{\rho_{i+1}(k) - \rho_i(k)}{\rho_i(k) + \kappa} - \frac{\delta T v_i(k) h_i(k)}{L_i \lambda_i [\rho_i(k) + \kappa]}, \quad (14)$$

$$H[\rho_i(k)] = \min \left\{ v_f \cdot \exp \left[ -\frac{1}{\alpha_i} \left( \frac{\rho_i(k)}{\rho_{cr}(k)} \right)^{\alpha_i} \right], b_i(k) v_f \right\}, \quad (15)$$

$$h_i(k) = \min r_i(k) \left\{ \left( d_i(k) + \frac{\omega_i(k)}{T} \right), Q_i(k) \frac{\rho_{jam} - \rho_i(k)}{\rho_{jam} - \rho_{cr}} \right\}, \quad (16)$$

$$\omega_i(k) = \omega_i(k-1) + T [d_i(k-1) - h_i(k-1)]. \quad (17)$$

The meanings and expressions of the parameters in the formula are the same as in Section 3.2.1, Section 3.2.2, and Section 3.2.3.

3.3. *Objective Function and Constraints.* In order to reflect the effect of ramp control, the ramp queuing time should be as small as possible, and the queuing length should be as short as possible. The total travel time of vehicles through the

where  $\delta$  is model parameter,  $\delta = 1 - v_{r,i}(k)/v_i(k)$ .

(2) *Ramp Inflow.* When RM is adopted, the traffic volume of the ramp merging into the mainline will also be affected by the on-ramp control rate. Therefore, the ramp control rate can be introduced as a coefficient when calculating the ramp merging volume, that is,

$$h_i(k) = \min r_i(k) \left\{ \left( d_i(k) + \frac{\omega_i(k)}{T} \right), Q_i(k) \frac{\rho_{jam} - \rho_i(k)}{\rho_{jam} - \rho_{cr}} \right\}, \quad (11)$$

where  $r_i(k)$  is on-ramp regulation rate.  $r_{\min,i} \leq r_i(k) \leq r_{\max,i}$  and  $r_{\min,i} \in [0, 1]$ .

3.2.4. *Establishment of Improved METANET Model.* The complete expression of the improved METANET model is as follows:

expressway and ramp can be divided into the sum of the time of passing the expressway mainline and the queuing time of the on-ramp. At the same time, the variable speed limit and ramp control should be as smooth and stable as possible to avoid large fluctuations and ensure road traffic safety. Therefore, the objective function and constraint conditions of the selected cooperative control model are described as follows:

$$\min J = T \sum_{j=1}^{N_p} \sum_{i=1}^M \lambda_i L_i [\rho_i(k+j)] + T \sum_{j=1}^{N_p} \omega(k+j) + T \sum_{j=1}^{N_p} \sum_{i=1}^M [r_i(k+j) - r_i(k+j-1)]^2 + T \sum_{j=1}^{N_p} \sum_{i=1}^M [b_i(k+j) - b_i(k+j-1)]^2, \quad (18)$$

$$\text{s.t.} \begin{cases} q_i(k) \leq Q_{\text{cap},i}(k) \\ |v_{i-1}(k) - v_i(k)| \leq 10 \\ |v_i(k) - v_i(k-1)| \leq 10, \\ \omega_i(k) \leq L_{\text{ramp}}, \end{cases} \quad (19)$$

where  $T$  is the sampling period,  $N_p$  is the calculation step,  $M$  is the number of sections, and  $Q_{cap,i}(k)$  is the capacity of each section of the mainline; the meaning of other parameters is the same as Section 3.2.

**3.4. Solution of Cooperative Control Model Based on NSGA-II Algorithm.** The essence of solving the cooperative control model is to solve the multiobjective optimization problem. Generally, the multiobjective optimization problem could not find a unique optimal solution but can find a set composed of multiple optimal solutions, that is, the Pareto optimal solution set. The conventional solution method is to convert it into a single objective problem for calculation. However, this has obvious defects that the weight selection of each objective depends on experience and the transformed problem is likely to be more complex, which has great limitations. Therefore, this paper selects the improved algorithm NSGA-II of nondominated sorting genetic algorithm (NSGA) to solve the proposed cooperative control model. Compared with other algorithms, this algorithm has a great improvement in performance. NSGA-II Algorithm 1 is specifically described as follows:

## 4. Example Simulation

In order to verify the effectiveness of the proposed collaborative control strategy, VISSIM simulation software is selected to simulate the traffic flow's operating conditions. VISSIM is a microscopic road traffic modeling tool that evaluates plans by simulating traffic conditions, taking full account of the traffic characteristics of motor vehicles, pedestrians, and so on. Its core simulation models include the following: (1) vehicle longitudinal movement adopts the psychophysical vehicle following model proposed by Professor Wiedemann; (2) vehicle lateral movement adopts the psychology-based Sparmann lane change behavior model; (3) VISSIM uses dynamic traffic allocation method in route selection [26, 27].

The simulation road network in this article comes from the Suzhou Central Elevated Expressway. The mainline of this expressway is about 2 km in length. It is a one-way three-lane lane. It is widened to four lanes at the entrance ramp and then becomes a three-lane lane with a lane width of 3.5 m. According to the requirements of the METANET model for road segment division, the research road segment is divided into 4 basic road segments, each with a length of 500 m. The simplified schematic diagram is shown in Figure 4.

This paper selects the evening peak period of 16:30–18:30 on a working day for research. The time interval of data acquisition is 20 s. The types of data collected include induction coil number, acquisition time, lane number, traffic flow, speed, occupancy, and so on. Through certain data processing, the collected data can meet the control requirements under the intelligent network environment so as to improve the accuracy of traffic state data. Since the traffic flow state of expressway is usually studied based on section data, it is necessary to integrate the collected lane data into

section data. The section data can be calculated by using the method in Table 1.

In Table 1,  $x$  is the lane number;  $n$  is the number of section lanes; OCC is the occupation ratio;  $L$  is the length of the detector;  $D$  is the vehicle length.

### 4.1. Model Parameter Calibration

**4.1.1. The Parameters of Improved METANET Model.** The parameters to be calibrated are  $(v_f, \rho_{cr}, \tau, \kappa, \delta, \alpha_i)$ . In this paper, based on peak hour data combined with least squares [28] quantitative calibration and the use of the nonlinear fitting function which is called Lsqnonlin in MATLAB software, the calibration results of each parameter are shown in Tables 2 and 3.

It can be seen from Table 2 that the road sections directly affected by the on-ramp of the expressway (such as road 2 and 3) have relatively greater impact on the actual traffic capacity of the road. The critical density of the road section is higher, and the free-flow speed is lower.

**4.1.2. The Parameters of NSGA-II Algorithm.** Four key parameters of NSGA-II solution algorithm need to be calibrated when solving collaborative control model. After experiments, the values of these parameters are shown in Table 4.

**4.2. Validation of Traffic Flow Simulation Model.** For validation of traffic flow simulation model, speed and density are selected as indicators. The calibrated parameter values are substituted into the METANET model. The numerical simulation function of traffic flow model is implemented by using MATLAB software. With this function, the flow and speed simulation results of each detector section in the study area are obtained and compared with the actual data. According to the characteristics of the road section, this article selects the mainline road 2 of the expressway as the object. Compare the actual value and the predicted result of the speed and density of the road section at each time in the simulation time under the same traffic volume, and use the image to show it, as shown in Figure 5.

According to the flow density equation in the traffic flow model, the traffic flow is equal to the speed multiplied by the density. The relationship between speed and density is complex. Therefore, the measured values of speed and density are compared with the predicted values in Figure 5. It can be seen from Figure 5 that the predicted density value of road 2 was basically consistent with the actual density value. The predicted speed value was 7.68% higher than the actual value. However, its change trend was very close to the change of the actual speed value, which reflected that the simulation effect was good. After calculation, the average absolute percentage error of the improved traffic flow model was 9.3%, so it can be concluded that the calibrated improved METANET model has good applicability.

Step 1: Randomly generate the initial population  $P_t$ ,  $t = 0$  whose size is  $N$  and set the number of iteration times and initial operation parameter values.  
 Step 2: Take the initial population as the parent population, select, cross, and mutate it, and get a new generation of offspring population  $Q_t$ ,  $t = 0$ .  
 Step 3: Combine  $P_t$  and  $Q_t$  to obtain a new population  $R_t$  with the scale of  $2N$ . The nondominated set  $F$  is obtained by fast nondominated sorting. The crowding degree of individuals in each layer is calculated, and the optimal  $N$  individuals are selected to form a new parent population  $P_{(t+1)}$  according to the nondominated sorting hierarchical number and crowding degree.  
 Step 4: Perform selection on population  $P_{(t+1)}$  and cross and mutate the selected individuals to obtain a new offspring population  $Q_{(t+1)}$ .  
 Step 5: Repeat the above operations and output result ending the operation until the number of iterations reaches the upper limit; otherwise return to step 2.

ALGORITHM 1: NSGA-II algorithm.

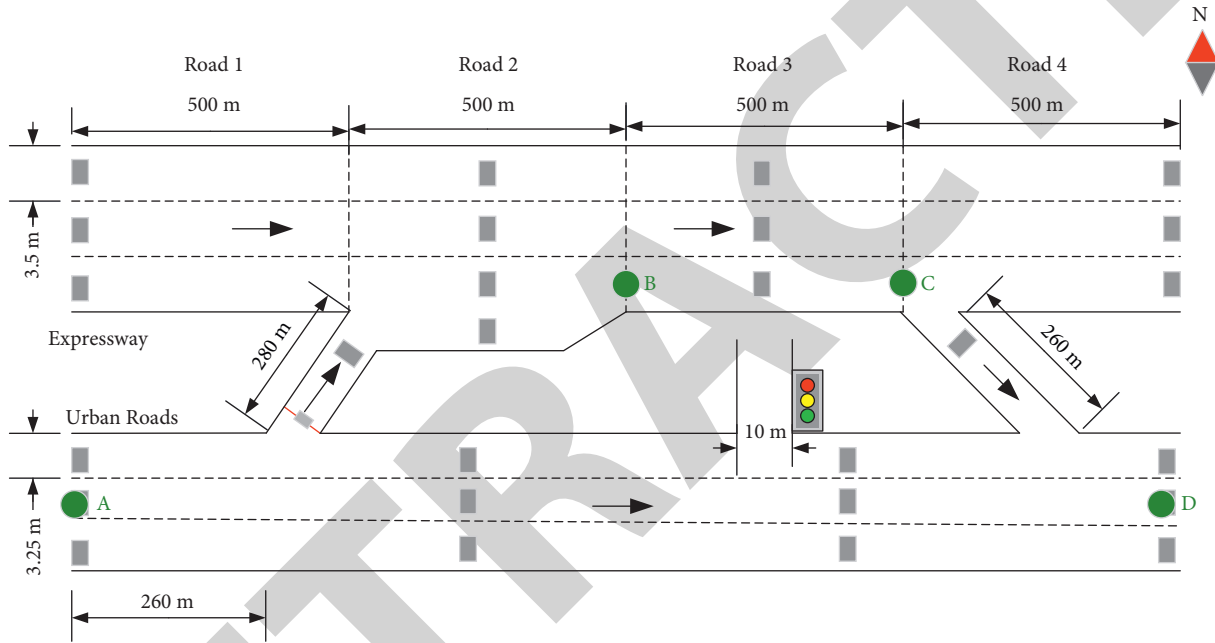


FIGURE 4: Schematic diagram of the application example.

TABLE 1: Traffic parameter collection method.

Data category	Calculation method
Sectional flow $q$	$q = \sum_{x=1}^n q_x$
Sectional velocity $v$	$v = \frac{\sum_{x=1}^n q_x v_x}{\sum_{x=1}^n q_x}$
Sectional density $\rho$	$\rho = 1/n \sum_{x=1}^n \text{occ}_x / (l + d)$

### 4.3. Analysis of Simulation Results

**4.3.1. Analysis of the Effect of the Expressway Ramp Cooperative Control Model.** Considering the actual traffic flow operating conditions in the study area, because there is an on-ramp in Section 2, which is in the expressway ramp confluence area, it is necessary to implement VSL on upstream Section 1 and implement RM on the on-ramp of Section 2. At the same time, it is necessary to analyse the queuing of vehicles in Section 2, so this article takes Section 2 as an example to analyse the effect of the control strategy.

It can be found from the left image of Figure 6 that when no control strategy is used, the traffic flow of road Section 2

encounters a bottleneck when the simulation time reaches about 39 minutes, and the traffic flow deceleration is very serious. This is because the average speed of the traffic flow decreases with the increase of the mainline vehicles and the on-ramp vehicles. When the simulation time reaches 65–90 min, the speed increase of traffic flow is not obvious when only using VSL. This is because the traffic flow of the mainline is too large at this time, and there is no control of the traffic flow on the ramp, resulting in insufficient lane changing gaps for the mainline vehicles, and the VSL control effect is very small. It can be clearly seen that when a serious traffic bottleneck is encountered, the expressway ramp coordinated control strategy has the highest benefit for the improvement of the mainline traffic flow speed.

It can be seen from the right picture of Figure 6, there is a significant increase in road section density during the first 30 minutes of the entire simulation period. This is due to the gradual increase in vehicles approaching the peak hours, and the density of road Section 2 increased due to the on-ramp vehicles merging. The use of variable speed limit control on

TABLE 2: Basic parameters of each road section.

Section number	Free-flow velocity $v_f$ (km/h)	Critical density $\rho_{cr}$ (veh/km/lane)
1	69.6	32.1
2	67.4	34.6
3	68.8	35.3
4	70.5	31.9

TABLE 3: METANET model parameters.

Parameter	$\tau$	$\kappa$	$\delta$	$\alpha_i$
Value	20	40	60	0.05

TABLE 4: NSGA-II Algorithm parameters.

Parameters	Size of population	Maximum algebra	Probability of cross	Probability of variation
Value	100	100	0.9	0.5

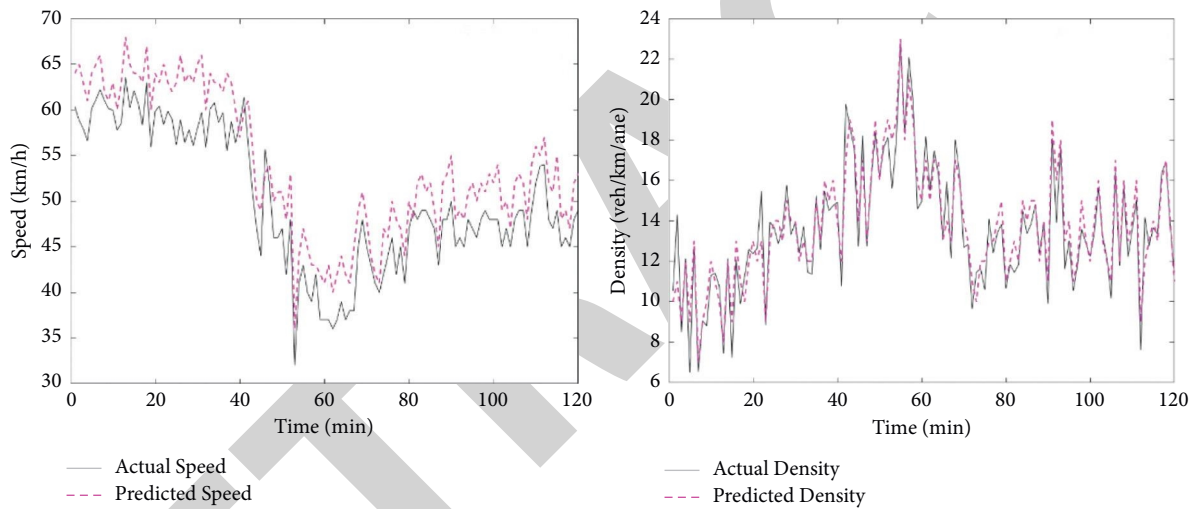


FIGURE 5: Comparison of speed and density prediction results of road 2.

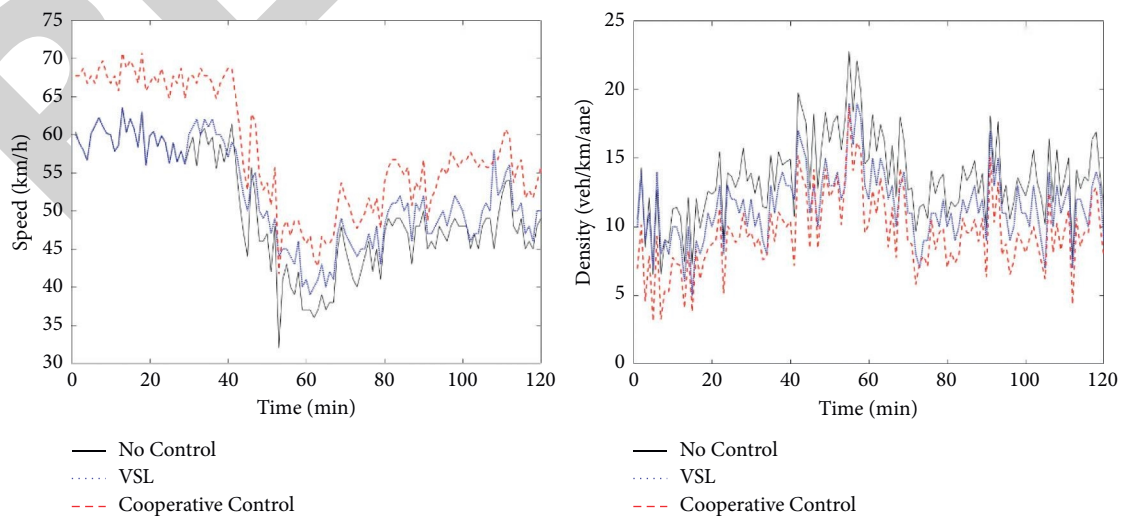


FIGURE 6: Speed and density changes of Section 2 under different control scenarios.



TABLE 5: Comparison of evaluation indicators under different control scenarios.

Evaluation index	No control	Cooperative control	Improve effect (%)	Joint control	Improve effect (%)
Total travel time/h	136	121	11.03	108	20.59
Mainline vehicle travel time/s	78	71	8.97	66	15.38
Average delay time/s	40.46	32.73	19.11	27.15	32.9
Speed difference/(km/h)	9.1	7.4	18.68	6.0	34.07

the upstream of the mainline (that is, Section 1) can reduce the number of vehicles entering Section 2 on the mainline, that is, reducing the upstream traffic demand. However, with the increase in traffic flow from the ramp, it is still impossible to change the congestion situation in the confluence area of the mainline and the on-ramp. After adopting the expressway ramp cooperative control, the average density of the road section is much smaller than before, which shows that the cooperative control of VSL and RM is very important.

#### 4.3.2. Comparative Analysis of Different Control Situations.

In order to quantitatively evaluate the effect of the proposed expressway ramp joint control system, this paper selects the total travel time, the mainline vehicle travel time, the average delay time, and the speed difference between adjacent road sections as evaluation indicators. Compare the operating efficiency and safety of the no-control strategy, only the expressway ramp cooperative control strategy, and the route decision control and expressway ramp cooperative control strategy. The comparison results are shown in Table 5.

It can be seen from the above table that, compared with the case of no control, the total travel time, mainline traffic travel time, average delay, and speed difference are reduced by 11.03%, 8.97%, 19.11%, and 18.68% after adopting the express ramp coordinated control strategy. With the addition of path decision control, the road network has been greatly improved in terms of operational efficiency and safety. In particular, the average delay time of the road network is reduced by 32.9% compared to the uncontrolled scenario, and the speed difference between adjacent road sections is reduced by 34.07%, and the operation of vehicles on the express road is more stable. Therefore, the effectiveness of the expressway ramp joint control system proposed in this paper can be verified. Compared with other control strategies, its control effect is generally better.

## 5. Conclusions

In this paper, a data-driven control strategy for urban express ramp is constructed. The improved METANET traffic flow model which can realize the cooperative control of VSL and RM is proposed and the genetic algorithm NSGA-II is designed to solve the optimal solution problem of the objective function and constraints aiming at traffic efficiency and traffic safety in the cooperative control model. Based on the existing VIS SIM microsimulation software and its COM interface, the ramp control function of expressway in the intelligent network environment is programmed and simulated. The results show that, compared with the traditional driving environment, the intelligent network environment as the premise can provide

more comprehensive traffic information for vehicles and realize real-time information interaction among vehicles, traffic information, and traffic environment on the road. This real-time and effective information exchange mode further ensures that the collaborative control strategy of expressway ramp was significantly better than the single RM or VSL. The total travel time and the speed difference between adjacent sections of expressway mainline are reduced by 20.59% and 34.07%, respectively, which improves the operation efficiency and safety of expressway network. The collaborative control strategy proposed in this paper reflects the synergy, real time, and accuracy in ramp control, which has certain enlightenment and reference significance for the transformation of expressway traffic control mode. On the one hand, the problem of traffic data lag would be solved faced by expressway traffic control for a long time with the support of intelligent network technology. On the other hand, its control object has changed from macrotraffic flow to microvehicles. The further improvement of control accuracy is more conducive to the improvement of traffic efficiency and safety.

Due to the limitation of time and ability, there are many deficiencies and parts to be improved in the research. In the future, the research can be further improved from the following aspects:

- (1) The optimization objective of the expressway ramp collaborative control model proposed in this paper is only considered from the perspective of traffic efficiency and safety, which is not comprehensive enough and fails to achieve the optimal system and user. The optimal system optimally considers the overall operation level of urban expressway traffic, including minimum overall delay and maximum overall throughput. The optimal user considers the priority of special vehicles, such as ambulances, fire engines, public transportation, etc. [29]. These control objectives need to be further studied and realized.
- (2) This paper describes the expressway traffic flow state through the macrotraffic flow model. When establishing the simulation model, the driving behavior was characterized only by modifying the parameters in the intelligent network environment. The consideration of the driver behavior was not accurate and comprehensive, which is possible to result in the difference between the simulation effect of the collaborative control model and the practical application. Therefore, the next research could consider using a more micromodel to describe the driving behavior in the intelligent network environment.



## Data Availability

The data were generated from experiments and can be made available by the corresponding author upon request.

## Conflicts of Interest

The authors declare no conflicts of interest.

## Acknowledgments

They would also like to acknowledge the support of the participating organization and its personnel who provided various effective assistance, such as data collection and the case study project. Lastly, they sincerely thank the financial support provided by the abovementioned mechanism. This research was funded by Postgraduate Research & Practice Innovation Program of Jiangsu Province (SJCX20\_1117, SJCX21\_1420, and KYCX21\_2999), Construction System Project of Jiangsu Province (2020ZD14 and 2018ZD258), Suzhou Social Science Fund (Y2020LX017 and Y2020LX025), General Project of Jiangsu Natural Fund (BK20151201 and BK20160357), and Philosophy and Social Science Projects of Universities in Jiangsu Province (2018SJA1348).

## References

- [1] J. Ma, L.-Y. Zhang, and K.-P. Li, "Traffic flow modeling and simulation of urban expressway and its ramp," *Computer Simulation*, vol. 33, no. 12, pp. 157–162+197, 2016.
- [2] Y.-W. Wang, "The research of iterative learning control method for speedway on-ramp," M.S. thesis, Chang'an University, Xi'an, China, 2008.
- [3] R. C. Carlson, I. Papamichail, and M. Papageorgiou, "Local feedback-based mainstream traffic flow control on motorways using variable speed limits," *IEEE Transactions on Intelligent Transportation Systems*, IEEE, vol. 12, no. 4, pp. 1261–1276, 2011.
- [4] E. Walraven, M. T. J. Spaan, and B. Bakker, "Traffic flow optimization: a reinforcement learning approach," *Engineering Applications of Artificial Intelligence*, vol. 52, 2016.
- [5] R. C. Carlson, I. Papamichail, and M. Papageorgiou, "Integrated ramp metering and mainstream traffic flow control on freeways using variable speed limits," *Transportation Research Part C: Emerging Technologies*, vol. 46, pp. 209–221, 2014.
- [6] Y. Han, M. Ramezani, A. Hegyi, Y. Yuanb, and S. Hoogendoornb, "Hierarchical ramp metering in freeways: an aggregated modeling and control approach," *Transportation Research*, vol. 110, no. ., pp. 1–19, 2020.
- [7] D.-X. Yu, H. Liu, L. L. Zhen, X. Ma, X. Xing, and H. Zhang, "Variable speed limit control method for freeway bottleneck area," *Journal of Transportation Systems Engineering and Information Technology*, vol. 18, no. 3, pp. 120–125, 2018.
- [8] Z.-Y. Lian, "Research on coordinated control strategy of on ramp and variable speed limit on Expressway," M.S. thesis, Harbin Institute of Technology, Harbin, China, 2016.
- [9] Z.-H. Qian, C.-S. Tian, Y.-J. Guo, and X. Wang, "The key technology and development of intelligent and connected transportation system," *Journal of Electronics and Information Technology*, vol. 42, no. 1, pp. 7–24, 2020.
- [10] K.-Q. Li, Y.-F. Dai, S.-B. Li, and M.-Y. Bian, "State-of-the-art and technical trends of intelligent and connected vehicles," *Journal of Automotive Safety and Energy*, vol. 8, 2017.
- [11] J.-Q. Wang and X. Wang, "Study on the system framework and key technology of Intelligent connected vehicles," *Journal of Chang'an University (Natural Science Edition)*, vol. 19, no. 6, pp. 18–25, 2017.
- [12] L. Qiao, H. Bao, Z.-X. Xuan, J. Liang, and F. Pan, "Autonomous driving ramp merging model based on reinforcement learning," *Computer Engineering*, vol. 44, no. 7, pp. 20–24, 2018.
- [13] C.-B. Zhang, J.-S. Li, C.-M. Huang, and Y.-X. Xia, "Freeway on-ramp vehicle merging guidance method based on vehicle-road coordination," *Journal of Wuhan University of Technology*, vol. 2017, no. 41, p. 542, 2017.
- [14] X.-L. Luo and Y.-S. Jiang, "Vehicle trajectory optimization model for ramp based on connect automatically vehicles," *Journal of Transportation Systems Engineering and Information Technology*, vol. 19, no. 4, pp. 94–100, 2019.
- [15] X. Miao, *Coordinated Ramp Metering under the Connected Vehicle Environment*, M.S. thesis, Southeast University, Nanjing, China, 2014.
- [16] Y.-L. Hu, Y.-F. Sun, and B.-C. Yin, "Information sensing and interaction technology in internet of things," *Chinese Journal of Computers*, vol. 35, no. 6, pp. 1147–1163, 2012.
- [17] F.-T. Li, "A brief analysis of the communication technology of the internet of vehicles," *Digital Communication World*, vol. 2020, no. 2, p. 87, 2020.
- [18] J. Ma, L.-Y. Zhang, and L.-X. Yan, "Macroscopic traffic flow: a modeling and simulation study considering impedance characteristics," *Advances in Transportation Studies*, vol. 2, pp. 162–174, 2019.
- [19] L.-Y. Zhang, J. Ma, B. Ran, and L.-X. Yan, "Traffic multi-resolution modelling and consistency analysis of urban expressway based on asynchronous integration strategy," *Modelling and Simulation in Engineering*, vol. 2017, Article ID 3694791, 19 pages, 2017.
- [20] C.-F. Shao, L.-Y. Wei, and B. Jiao, *The Theory of Traffic Flow*, Electronic Industry Press, Beijing, China, 2012.
- [21] M. J. Lighthill and G. B. Whitham, "On kinematic waves. II. A theory of traffic flow on long crowded roads," *Proceedings of The Royal Society A Mathematical Physical and Engineering Sciences*, vol. 229, no. 1178, pp. 317–345, 1955.
- [22] C. F. Daganzo, "The cell transmission model [II]: network traffic," *Transportation Research Part B: Methodological*, vol. 29, 1995.
- [23] A. Messmer and M. Papageorgiou, "METANET: a macroscopic simulation program for motorway networks," *Traffic Engineering and Control*, vol. 31, no. 8, pp. 466–470, 1990.
- [24] A. Kotsialos, M. Papageorgiou, C. Diakaki, Y. Pavlis, and F. Middelham, "Traffic flow modeling of large-scale motorway networks using the macroscopic modeling tool METANET," *IEEE Transactions on Intelligent Transportation Systems*, vol. 3, no. 4, pp. 282–292, 2002.
- [25] H. Zhou, J.-M. Hu, Y. Zhang, and R. Sun, "A coordinated optimization strategy of variable speed limit and ramp metering for expressway," *Journal of Transportation Systems Engineering and Information Technology*, vol. 17, no. 2, pp. 68–75, 2017.
- [26] PTV GROUP, *VISSIM 5.00 User's Manual*, Tech. Rep., PTVAG, Karlsruhe, Germany, 2008.
- [27] L.-D. Zhang, Y.-L. Wang, L. Jia, and J.-S. Pan, "A review of traffic simulation," *Computer Simulation*, vol. 23, no. 6, pp. 255–258, 2006.

## Research Article

# On the Generalized Decreasing Mean Time to Failure or Replaced Ordering

Haiyan Wang <sup>1</sup>, Diantong Kang <sup>2</sup>, and Lei Yan <sup>1</sup>

<sup>1</sup>Business School, Zhejiang Wanli University, Ningbo, Zhejiang 315100, China

<sup>2</sup>School of Mathematics and Statistics, Hexi University, Zhangye, Gansu 734000, China

Correspondence should be addressed to Lei Yan; [yanlei@zwu.edu.cn](mailto:yanlei@zwu.edu.cn)

Received 16 September 2021; Accepted 8 October 2021; Published 15 November 2021

Academic Editor: Daqing Gong

Copyright © 2021 Haiyan Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, we establish two new stochastic orders, DMTFR (decreasing mean time to failure or replaced) and GDMTFR (generalized decreasing mean time to failure or replaced), and mainly investigate properties of the GDMTFR order. Some characterizations of the GDMTFR order are given. The implication relationships between the DMTFR and the GDMTFR orders are considered. Also, closure and reversed closure properties of the new order GDMTFR are investigated. Meanwhile, several illustrative examples that meet the GDMTFR order are shown as well.

## 1. Introduction and Preliminaries

In risk investment fields, we often need to compare or select two risk assets [1]. Similarly, two lifetimes of two systems or units also need to be compared in engineering technologies [2]. To do these things, some refined stochastic orders were defined in statistics. For more details on stochastic orders, one may refer to the studies by Maria Fernandez-Ponce et al. [3], Belzunce [4], Müller and Stoyan [5], Li and Yam [6], Shaked and Shanthikumar [7, 8], Zhao and Balakrishnan [9], Sunoj et al. [10], Kang [11, 12], Yan [13], Kang and Yan [14], Vineshkumar [15], and the references therein. However, sometimes, we need to compare two risk assets by the aid of the third referred system, such as when we compare two risk assets [16], their values are changeable with the settlement time or with the kind of valuation currency [17]. To solve this problem, we establish a new stochastic order by introducing a common measure factor, namely, referred function.

Let  $X$  be a nonnegative and absolutely continuous random variable.  $X$  has distribution function  $F_X$  with 0 being the left endpoint of its support, survival function  $\bar{F}_X = 1 - F_X$ , and density function  $f_X$ . Let again  $F_X^{-1}$  be the right-continuous inverse function of  $F_X$  defined by

$$F_X^{-1}(p) = \sup\{x \mid F_X(x) \leq p\}, \quad 0 < p < 1. \quad (1)$$

In many important real areas, such as reliability, economics, management sciences, information sciences, and other related fields, stochastic comparisons are paid much attention and have been used as sharp tools in dealing with some random problems in recent years. Based on comparisons of residual life, mean residual life, mean inactivity time, failure rate, and many aging concepts are presented earlier or later. In these aspects, interested readers can refer to the studies by Barlow and Proschan [18], Shaked et al. [7], Shaked and Shanthikumar [8], Müller and Stoyan [5], Kochar et al. [19], Ahmad et al. [20], Knopik [21], Kochar et al. [19], and Li and Shaked [22] for some existing results.

In reality, a nonnegative random variable  $X$  is often called a life representing the lifetimes of a device. The life distribution classes IFR (increasing failure rate), IFRA (increasing failure rate average), DMRL (decreasing mean residual life), NBUE (new better than used in expectation), and MTFR (mean time to failure or replaced) classes of life distributions are commonly used to describe the ageing features of units or systems, for example, see Barlow and Proschan [18].

In particular, it is well known that (see Barlow and Campo [23], Klefsjö [24], Marschall [25])

$$\text{IFR} \subset \text{DMRL} \subset \text{NBUE}. \quad (2)$$

Knopik [21] introduced a new ageing class MTFR such that

$$\text{IFR} \subset \text{MTFR} \subset \text{NBUE}. \quad (3)$$

It is worthwhile to mention that, in Barlow's study [26], it is proved for the absolutely continuous distribution that

$$\text{IFRA} \subset \text{MTFR}. \quad (4)$$

Kochar et al. [19] showed that, for any random variable, the inclusion (4) is also valid.

Consider an age replacement policy as the one in which a unit is replaced by  $t$  time units after installation or at failure, whichever occurs first, and then the expected value for the first time to an in-service is (see Barlow and Proschan [18])

$$M_X(t) = \frac{\int_0^t \bar{F}_X(x) dx}{F_X(t)}, \quad \text{for } t \in \{x: F_X(x) > 0\}. \quad (5)$$

When  $M_X(t)$  is monotonic, the case was considered by Barlow and Campo [23], Marschall and Proschan [27], and Klefsjö [24]. In Knopik [21], and the ageing class MTFR (mean time to failure or replaced) of lifetime distribution is introduced, and it is proved that the MTFR class is closed under the operation of maximum for independence. In Kochar et al. [19], it is showed that the class MTFR is closed under weak convergence of distributions and convolution, and the dual family  $\text{MTFR}^D$  is closed under noncrossing mixtures.

*Definition 1.* The random variable  $X$  belongs to the mean time to failure with replacement (MTFR) class if the function  $X$  is decreasing for  $t \in \{x | F_X(x) > 0\}$ .

Assume that  $X$  has finite mean  $\mathbb{E}(X) = \mu_X$ . The residual life of  $X$  at time  $t > 0$  is defined as  $X_t = [X - t | X > t]$ . Then, the mean residual life of  $X$  at time  $t > 0$  is

$$\mu_X(t) = \mathbb{E}(X_t) = \frac{\int_t^{+\infty} \bar{F}_X(x) dx}{\bar{F}_X(t)}, \quad \text{for } t \geq 0. \quad (6)$$

If  $\mu_X(t)$  is decreasing, then we say that  $X$  (or  $F_X$ ) is in the decreasing mean residual (DMRL) life distribution class and denoted by  $X$  (or  $F_X$ )  $\in$  DMRL.

Emad-Edlin [28] proposed and studied the decreasing mean residual life (DMRL) ordering. Maria Fernandez-Ponce et al. [3] and Belzunce [4] proposed and studied the right spread order. Subsequently, Kochar et al. [19] proposed and studied the total time on test transform order and excess wealth order which is equivalent to the right spread order.

The total time on test (TTT) transform functions of  $X$  and  $Y$  are defined, respectively, as

$$\begin{aligned} T_X(p) &\equiv \int_0^{F_X^{-1}(p)} \bar{F}_X(x) dx, \\ T_Y(p) &\equiv \int_0^{G_Y^{-1}(p)} \bar{G}_Y(x) dx, \quad \text{for all } p \in (0, 1). \end{aligned} \quad (7)$$

The right spread (RS, for short) functions of  $X$  and  $Y$  are defined as, for all  $p \in (0, 1)$ , respectively,

$$\begin{aligned} S_X(p) &= \int_{F_X^{-1}(p)}^{\infty} \bar{F}_X(x) dx = \mathbb{E}[(X - F_X^{-1}(p))_+] = \mathbb{E}[\max\{X - F_X^{-1}(p), 0\}], \\ S_Y(p) &= \int_{G_Y^{-1}(p)}^{\infty} \bar{G}_Y(x) dx = \mathbb{E}[(Y - G_Y^{-1}(p))_+] = \mathbb{E}[\max\{Y - G_Y^{-1}(p), 0\}], \end{aligned} \quad (8)$$

where  $\bar{F}_X = 1 - F_X$  and  $\bar{G}_Y = 1 - G_Y$  are the survival functions of  $F_X$  and  $G_Y$ , respectively, and  $(x)_+ = \max\{x, 0\}$ .

Now, we recall several stochastic orders from Shaked and Shanthikumar [8].

*Definition 2.* Let  $X$  and  $Y$  be two nonnegative random variables with distribution functions  $F_X, G_Y$  such that  $F_X(0) = G_Y(0) = 0$ , survival functions  $\bar{F}_X \equiv 1 - F_X, \bar{G}_Y \equiv 1 - G_Y$ , right-continuous inverse functions  $F_X^{-1}, G_Y^{-1}$ , and finite means  $\mu_X, \mu_Y$ , respectively [8].

(1)  $X$  is said to be smaller than  $Y$  in the TTT transform order (denoted by  $X \leq_{\text{ttt}} Y$ ) if

$$\int_0^{F_X^{-1}(p)} \bar{F}_X(x) dx \leq \int_0^{G_Y^{-1}(p)} \bar{G}_Y(x) dx, \quad \text{for all } p \in (0, 1). \quad (9)$$

(2)  $X$  is said to be smaller than  $Y$  in the right spread order (denoted by  $X \leq_{\text{rs}} Y$ ) if

$$\int_{F_X^{-1}(p)}^{\infty} \bar{F}_X(x) dx \leq \int_{G_Y^{-1}(p)}^{\infty} \bar{G}_Y(x) dx, \quad \text{for all } p \in (0, 1). \quad (10)$$

(3)  $X$  is said to be smaller than  $Y$  in the NBUE (new better than used in expectation) order (denoted by  $X \leq_{\text{nbue}} Y$ ) if

$$\frac{\int_{F_X^{-1}(p)}^{\infty} \bar{F}_X(x) dx}{\int_{G_Y^{-1}(p)}^{\infty} \bar{G}_Y(x) dx} \leq \frac{\mu_X}{\mu_Y}, \quad \text{for all } p \in (0, 1). \quad (11)$$

(4)  $X$  is said to be smaller than  $Y$  in the location independent riskier order (denoted by  $X \leq_{\text{lir}} Y$ ) if

$$\int_0^{F_X^{-1}(p)} F_X(x) dx \leq \int_0^{G_Y^{-1}(p)} G_Y(x) dx, \quad \text{for all } p \in (0, 1). \quad (12)$$

(5)  $X$  is said to be smaller than  $Y$  in the DMRL order (denoted by  $X \leq_{\text{dmrl}} Y$ ) if

$$\frac{\mu_Y[G_Y^{-1}(p)]}{\mu_X[F_X^{-1}(p)]} = \frac{\int_{G_Y^{-1}(p)}^{\infty} \overline{G}_Y(x) dx}{\int_{F_X^{-1}(p)}^{\infty} \overline{F}_X(x) dx} \quad (13)$$

is increasing in  $p \in (0, 1)$ .

From Definition 1, we easily have the following result without proof.

**Lemma 1.** *Let  $X$  and  $Y$  be two continuous and nonnegative random variables with respective distribution functions  $F_X$  and  $G_Y$ , density functions  $f_X$  and  $g_Y$ , and the right-continuous inverse functions  $F_X^{-1}(p)$  and  $G_Y^{-1}(p)$  of  $F_X$  and  $G_Y$ , respectively. Then, the following statements are equivalent:*

- (1)  $X \leq_{\text{ttt}} Y$ .
- (2) *The inequality*

$$\int_0^{F_X^{-1}(p)} \overline{F}_X(x) dx \leq \int_0^{G_Y^{-1}(p)} \overline{G}_Y(x) dx, \quad \text{for all } p \in (0, 1), \quad (14)$$

is valid.

- (3) *The inequality*

$$\int_0^t \overline{F}_X(x) \left[ \frac{f_X(x)}{g_Y[G_Y^{-1}(F_X(x))]} - 1 \right] dx \geq 0, \quad \text{for all } t > 0, \quad (15)$$

is valid.

**Definition 3.** Let  $X$  and  $Y$  be two nonnegative random variables with distribution functions  $F_X, G_Y$  such that  $F_X(0) = G_Y(0) = 0$ , survival functions  $\overline{F}_X \equiv 1 - F_X, \overline{G}_Y \equiv 1 - G_Y$ , density functions  $f_X, g_Y$ , and right-continuous inverse functions  $F_X^{-1}, G_Y^{-1}$ , respectively.

- (1)  $X$  is said to be smaller than  $Y$  in the usual stochastic order (denoted by  $X \leq_{\text{st}} Y$ ), if  $\overline{F}_X(x) \leq \overline{G}_Y(x)$  for all  $x > 0$ .
- (2)  $X$  is said to be smaller than  $Y$  in the star-shaped order (denoted by  $X \leq_* Y$ ), if the function  $G_Y^{-1}(F_X(x))/x$  is increasing in  $x > 0$ .
- (3)  $X$  is said to be smaller than  $Y$  in the convex order (denoted by  $X \leq_c Y$ ), if the function  $G_Y^{-1}(F_X(x))$  is convex for  $x > 0$ .
- (4)  $X$  is said to be smaller than  $Y$  in the dispersive order (denoted by  $X \leq_{\text{disp}} Y$ ), if  $f_X[F_X^{-1}(p)] \geq g_Y[G_Y^{-1}(p)]$  for all  $p \in (0, 1)$ .

Let  $X$  and  $X$  be two absolutely continuous nonnegative random variables with respective distribution functions  $F_X$

and  $G_Y$  such that  $F_X(0) = G_Y(0) = 0$ , and density functions  $f_X$  and  $g_Y$ , and right-continuous inverse functions of  $F_X^{-1}$  and  $G_Y^{-1}$ , respectively.

We say that  $X$  is smaller than  $Y$  in the increasing concave order (denoted by  $X \leq_{\text{icv}} Y$ ) if

$$\int_0^t \overline{F}_X(x) dx \leq \int_0^t \overline{G}_Y(x) dx, \quad \text{for all } t > 0. \quad (16)$$

Here,

$$\begin{aligned} \int_0^t \overline{F}_X(x) dx &= \mathbb{E}[\min\{X, t\}], \\ \int_0^t \overline{G}_Y(x) dx &= \mathbb{E}[\min\{Y, t\}]. \end{aligned} \quad (17)$$

Hence, the order  $\leq_{\text{icv}}$  is a direct comparison for the mean service times of  $X$  and  $Y$ . In insurance theory,  $\int_0^d \overline{F}_X(x) dx = \mathbb{E}[\min\{X, d\}]$  is the average losses undertaken by the insured when the deductible excess is  $d$ . Hence,

$$\begin{aligned} T_X(p) &\equiv \int_0^{F_X^{-1}(p)} \overline{F}_X(x) dx = \mathbb{E}[\min\{X, F_X^{-1}(p)\}], \\ T_Y(p) &\equiv \int_0^{G_Y^{-1}(p)} \overline{G}_Y(x) dx = \mathbb{E}[\min\{Y, G_Y^{-1}(p)\}], \end{aligned} \quad (18)$$

that is,  $T_X(p)$  and  $T_Y(p)$  are the mean service times of  $X$  and  $Y$  before the respective equal probability time points. We wish to compare  $T_X(p)$  and  $T_Y(p)$  by means of a ratio  $T_Y(p)/T_X(p)$ . Based on such an idea, we give Definition 4.

Let  $X$  and  $Z$  be two nonnegative continuous random variables with respective distribution functions  $F_X$  and  $H_Z$  such that  $F_X(0) = H_Z(0) = 0$ , right-continuous inverse functions  $F_X^{-1}$  and  $H_Z^{-1}$ , survival functions  $\overline{F}_X$  and  $\overline{H}_Z$ , and density functions  $f_X$  and  $h_Z$ , respectively.

**Definition 4.**  $X$  is said to be smaller than  $Z$  in the decreasing mean time to failure (DMTFR) order (denoted by  $X \leq_{\text{dmtrf}} Z$ ), if

$$\frac{\int_0^{H_Z^{-1}(p)} \overline{H}_Z(x) dx}{\int_0^{F_X^{-1}(p)} \overline{F}_X(x) dx}, \quad (19)$$

is increasing in  $p \in (0, 1)$ .

From Definition 3, the following lemma is obvious, the proof is omitted here.

**Lemma 2.** *The following statements are equivalent:*

- (1)  $X \leq_{\text{dmtrf}} Z$ .
- (2) *The function  $T_Z(p)/T_X(p)$  is increasing in  $p \in (0, 1)$ .*
- (3) *The inequality*

$$\int_0^t \overline{F}_X(x) \left[ \frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} - \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \right] dx \geq 0, \quad \text{for all } t > 0, \quad (20)$$

is valid.

(4) The inequality

$$t\bar{F}_X(t)\left[\frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} - \frac{H_Z^{-1}(F_X(t))}{t}\right] - \int_0^t x\left(\frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} - \frac{H_Z^{-1}(F_X(x))}{x}\right)d[\bar{F}_X(x)] \geq 0, \quad \text{for all } t > 0, \quad (21)$$

is valid.

Shaked and Shanthikumar [8] studied the generalized TTT transform and proposed the generalized TTT transform order. Bartoszewicz and Benduch [29] further studied some properties of the generalized TTT transform by iteration. Motivated by their excellent works, we establish the following two new stochastic orders, the RDMTFR and the GDMTFR orders, see Definitions 5 and 6, respectively.

$$\int_0^t g_Y[G_Y^{-1}(F_X(x))]\left[\frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} - \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \right] dx \geq 0, \quad \text{for all } t > 0, \quad (23)$$

is valid.

(4) The inequality

$$tg_Y[G_Y^{-1}(F_X(t))]\left[\frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} - \frac{H_Z^{-1}(F_X(t))}{t}\right] - \int_0^t x\left(\frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} - \frac{H_Z^{-1}(F_X(x))}{x}\right)d[g_Y[G_Y^{-1}(F_X(x))]] \geq 0, \quad \text{for all } t > 0, \quad (24)$$

is valid.

It can be easily seen that when  $Y$  is an exponential random variable with a rate  $\lambda > 0$ , that is,  $Y \sim \mathbb{E}(\lambda)$ , there exists a simple relationship between the relative DM TFR order and the DM TFR order:

$$X \leq_{\text{rdmtfr}} Z \text{ r.t. } Y \iff X \leq_{\text{dmtfr}} Z. \quad (25)$$

Assume that  $\psi$  is a real function defined on  $(0, 1)$ . Denote by

$$T_X(p; \psi) = \int_0^{F_X^{-1}(p)} \psi[F_X(x)] dx, \\ T_Z(p; \psi) = \int_0^{H_Z^{-1}(p)} \psi[H_Z(x)] dx, \quad \text{for all } p \in (0, 1). \quad (26)$$

$T_X(p; \psi)$  and  $T_Z(p; \psi)$  are called the generalized TTT transforms of  $X$  and  $Z$  relative to  $\psi$ , respectively. Then, we give the following definition.

**Definition 5.** Let  $X, Z$ , and  $Y$  be three random variables.  $X$  is said to be smaller than  $Z$  relative to  $Y$  in the relative DM TFR order (denoted by  $X \leq_{\text{rdmtfr}} Z$  r.t.  $Y$ ), if

$$\frac{\int_0^{H_Z^{-1}(p)} g_Y[G_Y^{-1}(H_Z(x))] dx}{\int_0^{F_X^{-1}(p)} g_Y[G_Y^{-1}(F_X(x))] dx}, \quad (22)$$

is increasing in  $p \in (0, 1)$ .

From Definition 5, the following lemma is obvious, and the proof is omitted here.

**Lemma 3.** The following statements are equivalent:

- (1)  $X \leq_{\text{rdmtfr}} Z$  r.t.  $Y$ .
- (2) The function  $T_Z(p; Y)/T_X(p; Y)$  is increasing in  $p \in (0, 1)$ .
- (3) The inequality

**Definition 6.** Let  $X$  and  $Z$  be two absolutely continuous nonnegative random variables; assume that  $\psi$  is a real function defined on  $(0, 1)$ .  $X$  is said to be smaller than  $Z$  with respect to  $\psi$  in the generalized DM TFR order (denoted by  $X \leq_{\text{gdm tfr}} Z$  w.r.t.  $\psi$ ), if the function

$$\frac{\int_0^{H_Z^{-1}(p)} \psi[H_Z(x)] dx}{\int_0^{F_X^{-1}(p)} \psi[F_X(x)] dx}, \quad (27)$$

is increasing in  $p \in (0, 1)$ .

**Remark 1.** One can verify that the GDM TFR order is a partial order relation. The reflexive and transitive are evident; the antisymmetric is as follows (see Theorem 1 (30) as follows).

$X \leq_{\text{gdm tfr}} Z$  w.r.t.  $\psi$  and  $Z \leq_{\text{gdm tfr}} X$  w.r.t.  $\psi$  hold simultaneously, if and only if  $Y = kX + b$  almost surely, where  $k$  and  $b$  are any real numbers such that  $Y$  is nonnegative and  $k \neq 0$ .



*Remark 2.* It can be seen that the GDMTFR and DMRL orders do not have necessary implication relations each other. Especially, the DMTFR and DMRL orders do not have direct implication relationships each other.

From Definition 5, the following lemma is obvious, and the proof is omitted here.

It can be easily seen that when  $\psi(u) = g_Y[G_Y^{-1}(u)]$ ,  $u \in (0, 1)$ , there exists a simple relationship between the GDMTFR order and the RDMTFR order:

$$X \leq_{\text{gdmftr}} Z \text{ w.r.t. } \psi \iff X \leq_{\text{rdmftr}} Z \text{ r.t. } Y. \quad (28)$$

Now, we consider a system composed of same components. We say that the system preserves some properties if the components of this system possess some properties, and we conclude from the structure of the system that the system also possesses the same property. And conversely, we say that this system has the reversed preservation for some property if the system has some property. According to the structure of the system, we conclude that the components of this system also have the same property.

For two systems, if their two components satisfy some stochastic order relation, by the structure of these two systems, we conclude that the two systems also satisfy the same stochastic order relation; then, we say that these systems of the structure preserve this stochastic order relation, or equivalently, we say that this stochastic order relation possesses closure property under the structure. Conversely, if two systems satisfy some stochastic order relation, by the structure of these two systems, we conclude that the two components of the two systems also satisfy the same stochastic order relation, and then we say that these systems of the structure reversely preserve this stochastic order relation, or equivalently, we say that this stochastic order relation possesses reversed closure property under the structure.

In reliability theory, such two problems are often of interest: one is to investigate the closure properties of a stochastic order, and the other is to examine the reversed closure properties of a stochastic order under several reliability operations, such as increasing convex transforms and taking of maxima and minima.

To prove our main results, we first introduce the following lemma from Barlow and Proschan [18], which plays a key role in the proofs of this paper and are repeatedly used in the sequel.

**Lemma 4.** *Let  $W$  be a measure on the interval  $(a, b)$ , not necessarily nonnegative, where  $-\infty \leq a < b \leq +\infty$ . Let  $h$  be a nonnegative function defined on  $(a, b)$ . If  $\int_a^t dW(x) \geq 0$ , for all  $t \in (a, b)$  and if  $h$  is decreasing, then*

$$\int_a^b h(x) dW(x) \geq 0. \quad (29)$$

In the next, we assume that all the random variables, under consideration, are nonnegative and continuous.

Throughout this paper, the term *increasing* stands for monotone nondecreasing and *decreasing* stands for monotone nonincreasing. Assume that all the random variables under considerations are nonnegative and absolutely continuous with 0 as the left endpoint of their supports, and that all the integrals and expectations involved are always finite. All the encountered ratios are always supposed to be well defined.

In this paper, we devote our interest to the closure properties of the GDMTFR order. The paper is organized as follows. First, in Section 2, we consider characterizations of GDMTFR order. We investigate the implication relationships between DMTFR and GDMTFR orders in Section 3. The closure and reversed closure properties of the GDMTFR order are studied in Section 4. Finally, in Section 5, we give two examples which meet the GDMTFR order.

In the following, we always assume that  $X$  and  $Z$  are two absolutely continuous and nonnegative random variables with distribution functions  $F_X$  and  $H_Z$  such that  $F_X(0) = H_Z(0) = 0$ , survival functions  $\bar{F}_X \equiv 1 - F$  and  $\bar{H}_Z \equiv 1 - H_Z$ , density functions  $f_X$  and  $h_Z$ , and right-continuous inverse functions  $F_X^{-1}$  and  $H_Z^{-1}$  of  $F_X$  and  $H_Z$ , respectively.

## 2. Characterizations of the GDMTFR Order

Now, we explore some characterizations of the GDMTFR order. First, we give a result by Definition 1.6, which will be useful in the proofs of upcoming theorems.

**Theorem 1.** *The following statements are equivalent:*

- (1)  $X \leq_{\text{gdmftr}} Z$  w.r.t.  $\psi$ .
- (2) The inequality

$$\int_0^t \psi[F_X(x)] \left[ \frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} - \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \right] dx \geq 0, \quad \text{for all } t > 0, \quad (30)$$



is valid.

(3) The inequality

$$t\psi[F_X(t)]\left[\frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} - \frac{H_Z^{-1}(F_X(t))}{t}\right] - \int_0^t x\left(\frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} - \frac{H_Z^{-1}(F_X(x))}{x}\right)d[\psi(F_X(x))] \geq 0, \quad \text{for all } t > 0, \quad (31)$$

is valid.

$$\frac{\int_0^{H_Z^{-1}(p)} \psi[H_Z(x)]dx}{\int_0^{F_X^{-1}(p)} \psi[F_X(x)]dx}, \quad (32)$$

*Proof.* We only give proof for the case of (1) $\iff$ (2). Suppose that the function

is increasing in  $p \in (0, 1)$ . By differentiating, we have that the numerator of the derivative of this ratio is

$$\frac{\psi(p)}{h_Z[H_Z^{-1}(p)]} \int_0^{F_X^{-1}(p)} \psi[F_X(x)]dx - \frac{\psi(p)}{f_X[F_X^{-1}(p)]} \int_0^{H_Z^{-1}(p)} \psi[H_Z(x)]dx, \quad \text{for all } p \in (0, 1). \quad (33)$$

Since the function

is increasing in  $p \in (0, 1)$ , we have

$$\frac{\int_0^{H_Z^{-1}(p)} \psi[H_Z(x)]dx}{\int_0^{F_X^{-1}(p)} \psi[F_X(x)]dx}, \quad (34)$$

$$\int_0^{F_X^{-1}(p)} \frac{f_X[F_X^{-1}(p)]}{h_Z[H_Z^{-1}(p)]} \psi[F_X(x)]dx - \int_0^{H_Z^{-1}(p)} \psi[H_Z(x)]dx \geq 0, \quad \text{for all } p \in (0, 1). \quad (35)$$

Letting  $H_Z(x) = F_X(y)$  in the second integral of the left-hand side of inequality (35) and then letting  $F_X^{-1}(p) = t$ , we obtain that

$$\int_0^t \psi[F_X(x)]\left[\frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} - \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \right]dx \geq 0, \quad \text{for all } t > 0. \quad (36)$$

And, the above deduction is reversible. Therefore, the proof of the theorem is complete.  $\square$

*Definition 7.* Let  $X$  and  $Z$  be nonnegative absolutely continuous random variables with distribution functions  $F_X, H_Z$  such that  $F_X(0) = H_Z(0) = 0$ , density functions  $f_X, h_Z$ , and right-continuous inverse functions  $F_X^{-1}, H_Z^{-1}$ , respectively.

- (1)  $X$  is said to be smaller than  $Z$  in the starshaped order (denoted by  $X \leq_* Z$ ), if the function  $H_Z^{-1}(F_X(x))/x$  is increasing in  $x > 0$ .

- (2)  $X$  is said to be smaller than  $Z$  in the convex order (denoted by  $X \leq_{\mathcal{C}} Z$ ), if the function  $H_Z^{-1}(F_X(x))$  is convex for  $x > 0$ .

- (3)  $X$  is said to be smaller than  $Z$  in the dispersive order (denoted by  $X \leq_{\text{disp}} Z$ ), if  $f_X[F_X^{-1}(p)] \geq h_Z[H_Z^{-1}(p)]$  for all  $p \in (0, 1)$ .

**Theorem 2.** Let  $\psi$  be a nonnegative function defined on  $(0, 1)$ . Then, for any  $\theta > 0$ ,  $X \leq_{\text{gdmfr}} \theta X$  w.r.t.  $\psi$ .

**Theorem 3.** Let  $\psi$  be a nonnegative function defined on  $(0, 1)$ . If  $X \leq_{\text{gdmftr}} Z$  w.r.t.  $\psi$ , then for any  $\theta > 0$ ,  $X \leq_{\text{gdmftr}} \theta Z$  w.r.t.  $\psi$ .

*Proof.* Suppose that  $X \leq_{\text{gdmftr}} Z$  w.r.t.  $\psi$ , then from inequality (30), we have

$$\int_0^t \psi[F_X(x)] \left[ \frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} - \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \right] dx \geq 0, \quad \text{for all } t > 0. \quad (37)$$

In turn, from (30),  $X \leq_{\text{gdmftr}} \theta Z$  w.r.t.  $\psi$  if and only if, for all  $t > 0$ ,

$$\begin{aligned} & \int_0^t \psi[F_X(x)] \left[ \frac{f_X(t)}{h_{\theta Z}[H_{\theta Z}^{-1}(F_X(t))]} - \frac{f_X(x)}{h_{\theta Z}[H_{\theta Z}^{-1}(F_X(x))]} \right] dx \\ &= \theta \int_0^t \psi[F_X(x)] \left[ \frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} - \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \right] dx \geq 0. \end{aligned} \quad (38)$$

Clearly, if inequality (37) holds, then inequality (38) also holds, that is,  $X \leq_{\text{gdmftr}} \theta Z$  w.r.t.  $\psi$ . Therefore, the proof of the theorem is complete.  $\square$

*Remark 3.* Theorem 3 states that the order  $\leq_{\text{gdmftr}}$  is scale invariant with respect to the compared variables.

**Theorem 4.** If  $X \leq_{\text{gdmftr}} Z$  w.r.t.  $\psi$ , then for any  $a > 0$ ,  $X \leq_{\text{gdmftr}} aZ$  w.r.t.  $a\psi$ .

*Remark 4.* Theorem 4 states that the  $\leq_{\text{gdmftr}}$  order is scale invariant with respect to the referred function.

**Theorem 5.** If  $X \leq_{\text{gdmftr}} Z$  w.r.t.  $\psi$ , then for any  $\theta > 0$ ,  $\theta X \leq_{\text{gdmftr}} \theta Z$  w.r.t.  $\psi$ .

*Remark 5.* Theorem 5 states that the order  $\leq_{\text{gdmftr}}$  is scale equivalent with respect to the compared random variables.

**Theorem 6.** If  $X \leq_{\text{gdmftr}} Z$  w.r.t.  $\psi$ , then for any  $\theta > 0$  and  $a > 0$ ,  $X \leq_{\text{gdmftr}} \theta Z$  w.r.t.  $a\psi$ .

*Remark 6.* Theorem 6 states that the order  $\leq_{\text{gdmftr}}$  is scale invariant with respect to the compared random variables and the referred random variable.

**Theorem 7.** Let  $H_Z(0) = F_X(0) = 0$ .  $X \leq_{\text{gdmftr}} Z$  w.r.t.  $\psi$  and  $X \geq_{\text{gdmftr}} Z$  w.r.t.  $\psi$  if and only if there exists some  $\theta > 0$  such that  $F_X(x) = H_Z(\theta x)$ , that is,  $Z \stackrel{d}{=} \theta X$ , where " $\stackrel{d}{=}$ " means equality in distributions.

*Proof.* Suppose that  $X \leq_{\text{gdmftr}} Z$  w.r.t.  $\psi$  and  $X \geq_{\text{gdmftr}} Z$  w.r.t.  $\psi$ . From Definition 6, we have that there exists some  $\theta > 0$  such that

$$\frac{\int_0^{H_Z^{-1}(p)} \psi[H_Z(x)] dx}{\int_0^{F_X^{-1}(p)} \psi[F_X(x)] dx} = \theta. \quad (39)$$

That is,

$$\int_0^{H_Z^{-1}(p)} \psi[H_Z(x)] dx = \theta \int_0^{F_X^{-1}(p)} \psi[F_X(x)] dx. \quad (40)$$

Differentiating both sides of above equality, we obtain  $(H_Z^{-1}(p))' = (F_X^{-1}(p))'$ . Thus,  $H_Z^{-1}(p) = F_X^{-1}(p) + c$ , where  $c$  is any real number. By the assumption of  $H_Z(0) = F_X(0) = 0$ ,  $c = 0$  and then  $H_Z^{-1}(p) = F_X^{-1}(p)$ . Hence,  $Z \stackrel{d}{=} \theta X$ , and this is the stated result.

The following theorem gives a sufficient condition for the GDMTFR order.  $\square$

**Theorem 8.** Let  $\psi$  be a nonnegative function defined on  $(0, 1)$ . If  $X \leq_c Z$ , then  $X \leq_{\text{gdmftr}} Z$  w.r.t.  $\psi$ .

*Proof.* From (30)  $X \leq_{\text{gdmftr}} Z$  w.r.t.  $\psi$  if and only if

$$\int_0^t \psi[F_X(x)] \left[ \frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} - \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \right] dx \geq 0, \quad \text{for all } t > 0. \quad (41)$$

If  $X \leq_c Z$ , then the function  $f_X(x)/g_Y[G_Y^{-1}(F_X(x))]$  is increasing; hence,

$$\frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \leq \frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]}, \quad \text{for all } 0 \leq x \leq t. \quad (42)$$

Moreover,  $\psi[F_X(x)] \geq 0$  for all  $x \geq 0$ . From inequality (42), we see that inequality (41) holds. That is,  $X \leq_{\text{gdmftr}} Z$  w.r.t.  $\psi$ . Therefore, the proof of the theorem is complete.  $\square$

### 3. Implication Relationships between DMTFR and GDMTFR Orders

Recall from [5] that a function  $h: (0, 1) \rightarrow (0, 1)$  is called star-shaped (anti-star-shaped) if  $h(x)/x$  is increasing (decreasing).  $h$  is star-shaped (anti-star-shaped) with respect to the point  $(1, 1)$  if  $[1 - h(x)]/[1 - x]$  is increasing (decreasing) in  $x$ .

Now, we propose the following notions about real functions.

*Definition 8.* For a positive integer number  $n$ , we say that

- (1)  $h$  is star-shaped (anti-star-shaped) with respect to the point  $(1, 0)$  if  $h(x)/(1-x)$  is decreasing (increasing).
- (2)  $h$  is dual (anti-star-shaped) star-shaped of order  $n$  if  $h(x)/(1-x^n)$  is increasing (decreasing).
- (3)  $h$  is star-shaped (anti-star-shaped) of order  $n$  with respect to the point  $(1, 0)$  if  $h(x)/(1-x)^n$  is decreasing (increasing) in  $x \in (0, 1)$ .

Let  $X$  and  $Z$  be two nonnegative continuous random variables with respective distribution functions  $F_X$  and  $H_Z$  having 0 as the common left endpoint of their supports. Let again  $\psi$  be a nonnegative function defined on  $(0, 1)$ . Assume that  $X_1, \dots, X_n$  and  $Z_1, \dots, Z_n$  are the i.i.d. copies of  $X$  and  $Z$ , respectively, and that  $X_{k:n}$  and  $Z_{k:n}$ ,  $k = 1, 2, \dots, n$ , are the order statistics of  $X$  and  $Z$ , respectively. Let  $X_{k:n}$  have distribution function  $F_{X_{k:n}}(x)$ , survival function  $\bar{F}_{X_{k:n}}(x)$ , and density function  $f_{X_{k:n}}(x)$ , then for all  $x \geq 0$ ,

$$\begin{aligned} F_{X_{k:n}}(x) &= B_{k,n-k+1}(F_X(x)), \\ \bar{F}_{X_{k:n}}(x) &= B_{n-k+1,k}(\bar{F}_X(x)), \\ f_{X_{k:n}}(x) &= \beta_{k,n-k+1}(F_X(x))f_X(x) = \frac{n!}{(k-1)!(n-k)!} [F_X(x)]^{k-1} [\bar{F}_X(x)]^{n-k} f_X(x), \end{aligned} \quad (43)$$

where  $B_{k,n-k+1}(\cdot)$  and  $\beta_{k,n-k+1}(\cdot)$  are the distribution function and density function of a beta distribution with parameters  $k$  and  $n-k+1$ , respectively.

The following Theorem 9 gives some conditions under which the DMtFR order implies the GDMtFR order.

### Theorem 9

- (1) Assume that the function  $\psi \circ B_{k,n-k+1}$  is star-shaped with respect to the point  $(1, 0)$ . If  $X \leq_{\text{dmftr}} Z$ , then

$$X_{k:n} \leq_{\text{gdmftr}} Z_{k:n} \text{ w.r.t. } \psi, \quad \text{for all } k = 1, 2, \dots, n, \quad (44)$$

where  $\psi \circ B_{k,n-k+1}$  is the compound of  $\psi$  and  $B_{k,n-k+1}$ .

- (2) Assume that the function  $\psi$  is star-shaped with respect to the point  $(1, 0)$ . If  $X \leq_{\text{dmftr}} Z$ , then

$$X \leq_{\text{gdmftr}} Z \text{ w.r.t. } \psi. \quad (45)$$

- (3) Assume that the function  $\psi$  is star-shaped of order  $n$  with respect to the point  $(1, 0)$ . If  $X_{1:n} \leq_{\text{dmftr}} Z_{1:n}$  then

$$X \leq_{\text{gdmftr}} Z \text{ w.r.t. } \psi. \quad (46)$$

- (4) Assume that the function  $\psi$  is dual anti-star-shaped of order  $n$ . If  $X_{n:n} \leq_{\text{dmftr}} Z_{n:n}$  then

$$X \leq_{\text{gdmftr}} Z \text{ w.r.t. } \psi. \quad (47)$$

- (5) Assume that the function  $\psi$  star-shaped with respect to the point  $(1, 0)$ . If for some  $1 \leq k \leq n$ ,  $X_{k:n} \leq_{\text{dmftr}} Z_{k:n}$  then

$$X_{k:n} \leq_{\text{gdmftr}} Z_{k:n} \text{ w.r.t. } \psi, \quad \text{for all } k = 1, 2, \dots, n. \quad (48)$$

*Proof.* Barlow and Proschan [18] proved that, for all  $x \geq 0$ ,

$$H_{Z_{k:n}}^{-1}[F_{X_{k:n}}(x)] = \bar{H}_{Z_{k:n}}^{-1}[\bar{F}_{X_{k:n}}(x)] = H_Z^{-1}[F_X(x)]. \quad (49)$$

It can be proven that

$$\frac{f_{X_{k:n}}(x)}{h_{Z_{k:n}}[H_{Z_{k:n}}^{-1}(F_{X_{k:n}}(x))]} = \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]}. \quad (50)$$

- (1) Suppose that  $X \leq_{\text{dmftr}} Z$ . Then, we have, for all  $t > 0$ ,

$$\begin{aligned} & \int_0^t \bar{F}_X(x) \left[ \frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} - \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \right] dx \\ &= \int_0^t \bar{F}_X(x) \left[ \frac{f_{X_{k:n}}(t)}{h_{Z_{k:n}}[H_{Z_{k:n}}^{-1}(F_{X_{k:n}}(t))]} - \frac{f_{X_{k:n}}(x)}{h_{Z_{k:n}}[H_{Z_{k:n}}^{-1}(F_{X_{k:n}}(x))]} \right] dx \geq 0. \end{aligned} \quad (51)$$

Since  $\psi \circ B_{k,n-k+1}$  is star-shaped with respect to the point  $(1, 0)$ , then  $\psi \circ B_{k,n-k+1}(u)/(1-u)$  is decreasing in  $u \in (0, 1)$ . Hence, the function

$$h(x) = \psi \circ B_{k,n-k+1} \frac{[F_X(x)]}{(1-F_X(x))} = \psi \frac{[F_{X_{k:n}}(x)]}{\bar{F}_X(x)}, \quad (52)$$

is nonnegative decreasing in  $x \geq 0$ . From Lemma 4 and inequality (51), we have

$$\int_0^t \psi[F_{X_{k:n}}(x)] \left[ \frac{f_{X_{k:n}}(t)}{h_{Z_{k:n}}[H_{Z_{k:n}}^{-1}(F_{X_{k:n}}(t))]} - \frac{f_{X_{k:n}}(x)}{h_{Z_{k:n}}[H_{Z_{k:n}}^{-1}(F_{X_{k:n}}(x))]} \right] dx \geq 0, \quad \text{for all } t > 0, \quad (53)$$

which is equivalent to

$$X_{k:n} \leq_{\text{gdmftr}} Z_{k:n} \text{ w.r.t. } \psi, \quad \text{for all } k = 1, 2, \dots, n. \quad (54)$$

(2) Suppose that  $X \leq_{\text{dmftr}} Z$ . Then, for all  $t > 0$ ,

$$\int_0^t \bar{F}_X(x) \left[ \frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} - \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \right] dx \geq 0. \quad (55)$$

Since the function  $\psi$  is star-shaped with respect to the point  $(1, 0)$ , the function  $\psi(u)/(1-u)$  is decreasing in  $u \in (0, 1)$ . Thus,  $h(x) = \psi[F_X(x)]/\bar{F}_X(x)$  is nonnegatively decreasing. From Lemma 4 and (55), we have

$$\int_0^t \psi[F_X(x)] \left[ \frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} - \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \right] dx \geq 0, \quad \text{for all } t > 0, \quad (56)$$

which states that  $X \leq_{\text{gdmftr}} Z$  w.r.t.  $\psi$ .

(3) Suppose that  $X_{1:n} \leq_{\text{dmftr}} Z_{1:n}$ . Then, for all  $t > 0$ ,

$$\int_0^t \bar{F}_{X_{1:n}}(x) \left[ \frac{f_{X_{1:n}}(t)}{h_{Z_{1:n}}[H_{Z_{1:n}}^{-1}(F_{X_{1:n}}(t))]} - \frac{f_{X_{1:n}}(x)}{h_{Z_{1:n}}[H_{Z_{1:n}}^{-1}(F_{X_{1:n}}(x))]} \right] dx \geq 0. \quad (57)$$

Since  $\psi$  is star-shaped of order  $n$  with respect to the point  $(1, 0)$ , the function

$$\frac{\psi(x)}{(1-x)^n} \quad (58)$$

is decreasing in  $x$ , which leads to the function

$$h(x) = \frac{\psi[F_X(x)]}{[1-F_X(x)]^n} = \frac{\psi[F_X(x)]}{[\bar{F}_X(x)]^n} = \frac{\psi[F_X(x)]}{\bar{F}_{X_{1:n}}(x)}. \quad (59)$$

being nonnegative decreasing. Moreover,

$$\frac{f_{X_{1:n}}(x)}{h_{Z_{1:n}}[H_{Z_{1:n}}^{-1}(F_{X_{1:n}}(x))]} = \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \quad (60)$$

From Lemma 4, (57), and (60), we have

$$\int_0^t \psi[F_X(x)] \left[ \frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} - \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \right] dx \geq 0, \quad \text{for all } t > 0, \quad (61)$$

which is equivalent to

$$X \leq_{\text{gdmftr}} Z \text{ w.r.t. } \psi. \quad (62)$$

(4) Suppose that  $X_{n:n} \leq_{\text{dmftr}} Z_{n:n}$ . Then, for all  $t > 0$ ,

$$\int_0^t F_{X_{n:n}}(x) \left[ \frac{f_{X_{n:n}}(t)}{h_{Z_{n:n}}[H_{Z_{n:n}}^{-1}(F_{X_{n:n}}(t))]} - \frac{f_{X_{n:n}}(x)}{h_{Z_{n:n}}[H_{Z_{n:n}}^{-1}(F_{X_{n:n}}(x))]} \right] dx \geq 0. \quad (63)$$

Since  $\psi$  is dual anti-star-shaped of order  $n$ , the function  $\psi(x)/(1-x^n)$  is nonnegative decreasing in  $x \in (0, 1)$ , which leads to the function

$$h(x) = \frac{\psi[F_X(x)]}{1 - [F_X(x)]^n} = \frac{\psi[F_X(x)]}{\bar{F}_{X_{n:n}}(x)}. \quad (64)$$

being nonnegative decreasing. Moreover,

$$\frac{f_{X_{n:n}}(x)}{h_{Z_{n:n}}[H_{Z_{n:n}}^{-1}(F_{X_{n:n}}(x))]} = \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \quad (65)$$

From Lemma 4, (63)–(65), we have

$$\int_0^t \psi[F_X(x)] \left[ \frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} - \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \right] dx \geq 0, \quad \text{for all } t > 0, \quad (66)$$

and this asserts that

$$X \leq_{\text{gdmftr}} Z \text{ w.r.t. } \psi. \quad (67)$$

(5) Let the function  $\psi$  be anti-star-shaped with respect to the point  $(1, 0)$ . Suppose that, for some  $1 \leq k \leq n$ ,  $X_{k:n} \leq_{\text{dmftr}} Z_{k:n}$ , then for all  $t > 0$ ,

$$\int_0^t F_{X_{k:n}}(x) \left[ \frac{f_{X_{k:n}}(t)}{h_{Z_{k:n}}[H_{Z_{k:n}}^{-1}(F_{X_{k:n}}(t))]} - \frac{f_{X_{k:n}}(x)}{h_{Z_{k:n}}[H_{Z_{k:n}}^{-1}(F_{X_{k:n}}(x))]} \right] dx \geq 0. \quad (68)$$

Since the function  $\psi$  is star-shaped with respect to the point  $(1, 0)$ , the function  $\psi(x)/(1-x)$  is nonnegatively decreasing in  $x \in (0, 1)$ . Thus,  $\varphi(x) = \psi[F_{X_{k:n}}(x)]/\bar{F}_{X_{k:n}}(x)$

is nonnegatively decreasing. From Lemma 4 and (68), we have

$$\int_0^t \psi[F_{X_{k:n}}(x)] \left[ \frac{f_{X_{k:n}}(t)}{h_{Z_{k:n}}[H_{Z_{k:n}}^{-1}(F_{X_{k:n}}(t))]} - \frac{f_{X_{k:n}}(x)}{h_{Z_{k:n}}[H_{Z_{k:n}}^{-1}(F_{X_{k:n}}(x))]} \right] dx \geq 0, \quad \text{for all } t > 0, \quad (69)$$

which asserts that

$$X_{k:n} \leq_{\text{gdmftr}} Z_{k:n} \text{ w.r.t. } \psi, \quad \text{for all } k = 1, 2, \dots, n. \quad (70)$$

Therefore, the proof of the theorem is complete.

Theorem 10 gives some conditions under which the GDMTFR order implies the DMFTR order.  $\square$

### Theorem 10

(1) Assume that the function  $\psi$  is anti-star-shaped with respect to the point  $(1, 0)$ . If for some  $1 \leq k \leq n$ ,  $X_{k:n} \leq_{\text{gdmftr}} Z_{k:n}$  w.r.t.  $\psi$ , then

$$X_{k:n} \leq_{\text{dmftr}} Z_{k:n}, \quad \text{for all } k = 1, 2, \dots, n. \quad (71)$$

(2) Assume that the function  $\psi$  is anti-star-shaped with respect to the point  $(1, 0)$ . If  $X \leq_{\text{gdmftr}} Z$  w.r.t.  $\psi$ , then

$$X \leq_{\text{dmftr}} Z. \quad (72)$$

(3) Assume that the function  $\psi$  is dual star-shaped of order  $n$ . If  $X \leq_{\text{gdmftr}} Z$  w.r.t.  $\psi$ , then

$$X_{n:n} \leq_{\text{dmftr}} Z_{n:n}. \quad (73)$$

(4) Assume that the function  $\psi$  is anti-star-shaped of order  $n$  with respect to the point  $(1, 0)$ . If  $X \leq_{\text{gdmftr}} Z$  w.r.t.  $\psi$ , then

$$X_{1:n} \leq_{\text{dmtr}} Z_{1:n}. \quad (74) \quad \text{Proof}$$

(5) Assume that the function  $\psi$  is increasing in  $u \in (0, 1)$ .  
If  $X \leq_{\text{dmtr}} Z$  w.r.t.  $\psi$ , then

$$X_{k:n} \leq_{\text{gdmtr}} Z_{k:n}, \quad \text{for all } k = 1, 2, \dots, n. \quad (75)$$

(1) Suppose that, for some  $1 \leq k \leq n$ ,  $X_{k:n} \leq_{\text{gdmtr}} Z_{k:n}$   
w.r.t.  $\psi$ . Then, for all  $t > 0$ ,

$$\int_0^t \psi[F_{X_{k:n}}(x)] \left[ \frac{f_{X_{k:n}}(t)}{h_{Z_{k:n}}[H_{Z_{k:n}}^{-1}(F_{X_{k:n}}(t))]} - \frac{f_{X_{k:n}}(x)}{h_{Z_{k:n}}[H_{Z_{k:n}}^{-1}(F_{X_{k:n}}(x))]} \right] dx \geq 0. \quad (76)$$

Since the function  $\psi$  is anti-star-shaped with respect to the point  $(1, 0)$ , the function  $(1-u)/\psi(u)$  is nonnegatively decreasing in  $u \in (0, 1)$ . Thus,  $h(x) =$

$\bar{F}_{X_{k:n}}(x)/\psi[F_{X_{k:n}}(x)]$  is nonnegatively decreasing. From Lemma 4 and (76), we have

$$\int_0^t \bar{F}_{X_{k:n}}(x) \left[ \frac{f_{X_{k:n}}(t)}{h_{Z_{k:n}}[H_{Z_{k:n}}^{-1}(F_{X_{k:n}}(t))]} - \frac{f_{X_{k:n}}(x)}{h_{Z_{k:n}}[H_{Z_{k:n}}^{-1}(F_{X_{k:n}}(x))]} \right] dx \geq 0, \quad \text{for all } t > 0, \quad (77)$$

which asserts that

$$X_{k:n} \leq_{\text{dmtr}} Z_{k:n}, \quad \text{for all } k = 1, 2, \dots, n. \quad (78)$$

(2) Suppose that  $X \leq_{\text{gdmtr}} Z$  w.r.t.  $\psi$ . Then, for all  $t > 0$ ,

$$\int_0^t \psi[F_X(x)] \left[ \frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} - \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \right] dx \geq 0. \quad (79)$$

Since the function  $\psi$  is anti-star-shaped with respect to the point  $(1, 0)$ , the function  $(1-u)/\psi(u)$  is decreasing in  $u \in (0, 1)$ . Thus,  $h(x) = \bar{F}_X$

$(x)/\psi[F_X(x)]$  is nonnegatively decreasing. From Lemma 4 and (79), we have

$$\int_0^t \bar{F}_X(x) \left[ \frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} - \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \right] dx \geq 0, \quad \text{for all } t > 0, \quad (80)$$

which states that

$$X \leq_{\text{dmtr}} Z. \quad (81)$$

(3) Suppose that  $X \leq_{\text{gdmtr}} Z$  w.r.t.  $\psi$ . Then, for all  $t > 0$ ,

$$\int_0^t \psi[F_X(x)] \left[ \frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} - \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \right] dx \geq 0, \quad \text{for all } t > 0. \quad (82)$$

Since  $\psi$  is dual star-shaped of order  $n$ , the function  $(1-u^n)/\psi(u)$  is nonnegative decreasing in  $u \in (0, 1)$ , which leads to the function

$$\varphi(x) = \frac{1 - [F_X(x)]^n}{\psi[F_X(x)]} = \frac{\bar{F}_{X_{n:n}}(x)}{\psi[F_X(x)]}, \quad (83)$$

being nonnegative decreasing. Moreover,



$$\frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} = \frac{f_{X_{n:n}}(x)}{h_{Z_{n:n}}[H_{Z_{n:n}}^{-1}(F_{X_{1:n}}(x))]} \quad (84)$$

From Lemma 4, (82)–(84), we have

$$\int_0^t \bar{F}_{X_{n:n}}(x) \left[ \frac{f_{X_{n:n}}(t)}{h_{Z_{n:n}}[H_{Z_{n:n}}^{-1}(F_{X_{n:n}}(t))]} - \frac{f_{X_{n:n}}(x)}{h_{Z_{n:n}}[H_{Z_{n:n}}^{-1}(F_{X_{n:n}}(x))]} \right] dx \geq 0, \quad \text{for all } t > 0, \quad (85)$$

and this asserts that

$$X_{n:n} \leq_{\text{dmtr}} Z_{n:n}. \quad (86)$$

(4) Suppose that  $X \leq_{\text{gdmtr}} Z$  w.r.t.  $\psi$ . Then, for all  $t > 0$ ,

$$\int_0^t \psi[F_X(x)] \left[ \frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} - \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \right] dx \geq 0, \quad \text{for all } t > 0. \quad (87)$$

Since  $\psi$  is anti-star-shaped of order  $n$  with respect to the point  $(1, 0)$ , the function  $(1-u)^n/\psi(u)$  is decreasing in  $u \in (0, 1)$ , and this leads to the function

$$h(x) = \frac{[1 - F_X(x)]^n}{\psi[F_X(x)]} = \frac{\bar{F}_{X_{1:n}}(x)}{\psi[F_X(x)]}, \quad (88)$$

being nonnegatively decreasing. Moreover,

$$\frac{f_{X_{1:n}}(x)}{h_{Z_{1:n}}[H_{Z_{1:n}}^{-1}(F_{X_{1:n}}(x))]} = \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \quad (89)$$

From Lemma 4, (87)–(89), we have

$$\int_0^t \bar{F}_{X_{1:n}}(x) \left[ \frac{f_{X_{1:n}}(t)}{h_{Z_{1:n}}[H_{Z_{1:n}}^{-1}(F_{X_{1:n}}(t))]} - \frac{f_{X_{1:n}}(x)}{h_{Z_{1:n}}[H_{Z_{1:n}}^{-1}(F_{X_{1:n}}(x))]} \right] dx \geq 0, \quad \text{for all } t > 0, \quad (90)$$

which is equivalent to

$$X_{1:n} \leq_{\text{dmtr}} Z_{1:n}. \quad (91)$$

(5) Suppose that  $X \leq_{\text{dmtr}} Z$  w.r.t.  $\psi$ . Then, we have, for all  $t > 0$ ,

$$\begin{aligned} & \int_0^t \psi[F_X(x)] \left[ \frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} - \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \right] dx \\ &= \int_0^t \psi[F_X(x)] \left[ \frac{f_{X_{k:n}}(t)}{h_{Z_{k:n}}[H_{Z_{k:n}}^{-1}(F_{X_{k:n}}(t))]} - \frac{f_{X_{k:n}}(x)}{h_{Z_{k:n}}[H_{Z_{k:n}}^{-1}(F_{X_{k:n}}(x))]} \right] dx \geq 0. \end{aligned} \quad (92)$$

Since  $\psi$  is increasing in  $u \in (0, 1)$ , then  $\bar{F}_{X_{k:n}}(x)/\psi[F_X(x)]$  is nonnegatively decreasing in  $x > 0$ . From Lemma 4 and inequality (92), we have

$$\int_0^t \psi[F_{X_{k:n}}(x)] \left[ \frac{f_{X_{k:n}}(t)}{h_{Z_{k:n}}[H_{Z_{k:n}}^{-1}(F_{X_{k:n}}(t))]} - \frac{f_{X_{k:n}}(x)}{h_{Z_{k:n}}[H_{Z_{k:n}}^{-1}(F_{X_{k:n}}(x))]} \right] dx \geq 0, \quad \text{for all } t > 0, \quad (93)$$

which is equivalent to

$$X_{k:n} \leq_{\text{gdmftr}} Z_{k:n}, \quad \text{for all } k = 1, 2, \dots, n. \quad (94)$$

Therefore, the proof of the theorem is complete.

Denote by  $X_{1:N} = \min(X_1, X_2, \dots, X_N)$  and  $X_{N:N} = \max(X_1, X_2, \dots, X_N)$ .  $Z_{1:N}$  and  $Z_{N:N}$  are similar. Let  $\varphi_N(\cdot)$  be the probability generating function of  $N$ .  $\square$

**Theorem 11.** Assume that  $X_1, X_2, \dots$ , and  $Z_1, Z_2, \dots$ , are the i.i.d. copies of  $X$  and  $Z$ , respectively.

(1) Let the function  $\psi \circ \varphi_N$  be star-shaped with respect to the point  $(1, 0)$ . If  $X \leq_{\text{dmftr}} Z$ , then

$$X_{N:N} \leq_{\text{gdmftr}} Z_{N:N} \text{ w.r.t. } \psi. \quad (95)$$

(2) Let the function  $\psi \circ \varphi_N$  be anti-star-shaped with respect to the point  $(1, 0)$ . If  $X_{N:N} \leq_{\text{gdmftr}} Z_{N:N}$  w.r.t.  $\psi$ , then

$$X \leq_{\text{dmftr}} Z. \quad (96)$$

(3) Let  $\psi$  be a function defined on the interval  $(0, 1)$  such that the function  $\psi(1 - \varphi_N(u))/u$  is increasing in  $u \in (0, 1)$ . If  $X \leq_{\text{dmftr}} Z$ , then

$$X_{1:N} \leq_{\text{gdmftr}} Z_{1:N} \text{ w.r.t. } \psi. \quad (97)$$

(4) Let  $\psi$  be a function defined on the interval  $(0, 1)$  such that the function  $u/\psi(1 - \varphi_N(u))$  is increasing in  $u \in (0, 1)$ . If  $X_{1:N} \leq_{\text{gdmftr}} Z_{1:N}$  w.r.t.  $\psi$ , then

$$X \leq_{\text{dmftr}} Z. \quad (98)$$

*Proof*

(1) Suppose that  $X \leq_{\text{dmftr}} Y$ . This asserts that

$$\int_0^t \bar{F}_X(x) \left[ \frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} - \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \right] dx \geq 0, \quad \text{for all } t > 0. \quad (99)$$

Shaked and Shanthikumar [8] showed that

$$H_{Z_{N:N}}^{-1}(F_{X_{N:N}}(x)) = H_Z^{-1}(F_X(x)). \quad (100)$$

It can be proven that

$$\frac{f_{X_{N:N}}(x)}{h_{Z_{N:N}}[H_{Z_{N:N}}^{-1}(F_{X_{N:N}}(x))]} = \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \quad (101)$$

Since the function  $\psi \circ \varphi_N$  is star-shaped with respect to the point  $(1, 0)$ , the function  $\psi[\varphi_N(u)]/(1-u)$  is decreasing in  $u \in (0, 1)$ , and this leads to the function

$$h(x) = \frac{\psi[\varphi_N(F_X(x))]}{(1 - F_X(x))} = \frac{\psi[F_{X_{N:N}}(x)]}{\bar{F}_X(x)}, \quad (102)$$

being nonnegative decreasing. From Lemma 4, (99), and (101), we have for all  $t > 0$ ,

$$\int_0^t \psi[F_{X_{N:N}}(x)] \left[ \frac{f_{X_{N:N}}(t)}{h_{Z_{N:N}}[H_{Z_{N:N}}^{-1}(F_{X_{N:N}}(t))]} - \frac{f_{X_{N:N}}(x)}{h_{Z_{N:N}}[H_{Z_{N:N}}^{-1}(F_{X_{N:N}}(x))]} \right] dx \geq 0. \quad (103)$$

That is,  $X_{N:N} \leq_{\text{gdmftr}} Z_{N:N}$  w.r.t.  $\psi$ .

(2) Suppose that  $X_{N:N} \leq_{\text{gdmftr}} Z_{N:N}$  w.r.t.  $\psi$ . This asserts that, for all  $t > 0$ ,

$$\int_0^t \psi[F_{X_{N:N}}(x)] \left[ \frac{f_{X_{N:N}}(t)}{h_{Z_{N:N}}[H_{Z_{N:N}}^{-1}(F_{X_{N:N}}(t))]} - \frac{f_{X_{N:N}}(x)}{h_{Z_{N:N}}[H_{Z_{N:N}}^{-1}(F_{X_{N:N}}(x))]} \right] dx \geq 0. \quad (104)$$

Since the function  $\psi \circ \varphi_N(\cdot)$  is anti-star-shaped with respect to the point  $(1, 0)$ , the function  $(1-u)/\psi[\varphi_N(u)]$  is decreasing in  $u \in (0, 1)$ , and this leads to the function

$$h(x) = \frac{(1 - F_X(x))}{\psi[\varphi_N(F_X(x))]} = \frac{\bar{F}_X(x)}{\psi[F_{X_{N:N}}(x)]}, \quad (105)$$

being nonnegatively decreasing. From Lemma 4, (104), and (101), we have

$$\int_0^t \bar{F}_X(x) \left[ \frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} - \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \right] dx \geq 0, \quad \text{for all } t > 0. \quad (106)$$

That is,  $X \leq_{\text{dmftr}} Y$ .

(3) Suppose that  $X \leq_{\text{dmftr}} Z$ . Then,

$$\int_0^t \bar{F}_X(x) \left[ \frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} - \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \right] dx \geq 0, \quad \text{for all } t > 0. \quad (107)$$

Shaked and Shanthikumar [8] showed that

$$H_{Z_{1:N}}^{-1}(F_{X_{1:N}}(x)) = H_Z^{-1}(F_X(x)). \quad (108)$$

It can be proven that

$$\frac{f_{X_{1:N}}(x)}{h_{Z_{1:N}}[H_{Z_{1:N}}^{-1}(F_{X_{1:N}}(x))]} = \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \quad (109)$$

Since the function  $\psi(1 - \varphi_N(u))/u$  is increasing in  $u \in (0, 1)$ , the function

$$h(x) = \frac{\psi[\varphi_N(\bar{F}_X(x))]}{\bar{F}_X(x)} = \frac{\psi[F_{X_{1:N}}(x)]}{\bar{F}_X(x)}, \quad (110)$$

is nonnegative decreasing. From Lemma 4, (107), and (109), we have, for all  $t > 0$ ,

$$\int_0^t \psi[F_{X_{1:N}}(t)] \left[ \frac{f_{X_{1:N}}(t)}{h_{Z_{1:N}}[H_{Z_{1:N}}^{-1}(F_{X_{1:N}}(t))]} - \frac{f_{X_{1:N}}(x)}{h_{Z_{1:N}}[H_{Z_{1:N}}^{-1}(F_{X_{1:N}}(x))]} \right] dx \geq 0. \quad (111)$$

That is,  $X_{1:N} \leq_{\text{gdmftr}} Z_{1:N}$  w.r.t.  $\psi$ .

(4) Suppose that  $X_{1:N} \leq_{\text{gdmftr}} Z_{1:N}$  w.r.t.  $\psi$ . Then, for all  $t > 0$ ,

$$\int_0^t \psi[F_{X_{1:N}}(x)] \left[ \frac{f_{X_{1:N}}(t)}{h_{Z_{1:N}}[H_{Z_{1:N}}^{-1}(F_{X_{1:N}}(t))]} - \frac{f_{X_{1:N}}(x)}{h_{Z_{1:N}}[H_{Z_{1:N}}^{-1}(F_{X_{1:N}}(x))]} \right] dx \geq 0. \quad (112)$$

Since the function  $u/\psi(1 - \varphi_N(u))$  is increasing in  $u \in (0, 1)$ , the function

$$h(x) = \frac{\bar{F}_X(x)}{\psi[\varphi_N(\bar{F}_X(x))]} = \frac{\bar{F}_X(x)}{\psi[F_{X_{1:N}}(x)]}, \quad (113)$$

is nonnegatively decreasing. From Lemma 4, (112), and (109), we have for all  $t > 0$ ,

$$\int_0^t \bar{F}_X(x) \left[ \frac{f_{X_{1:N}}(t)}{h_{Z_{1:N}}[H_{Z_{1:N}}^{-1}(F_{X_{1:N}}(t))]} - \frac{f_{X_{1:N}}(x)}{h_{Z_{1:N}}[H_{Z_{1:N}}^{-1}(F_{X_{1:N}}(x))]} \right] dx \geq 0. \quad (114)$$

That is,  $X_{1:N} \leq_{\text{gdmftr}} Z_{1:N}$ . Therefore, the proof of the theorem is complete.  $\square$

**Theorem 12.** Assume that  $X \leq_{\text{gdmftr}} Z$  w.r.t.  $\psi$ ,  $\psi(u) = B_{n-k+1,k}(u)$ ,  $u \in (0, 1)$ , for some  $1 \leq k \leq n$ , where

$B_{n-k+1,k}$  is the distribution function of a beta distribution with parameters  $n - k + 1$  and  $k$ . Then, we have

$$X_{k:n} \leq_{\text{dmtfr}} Z_{k:n}, \quad \text{for all } k = 1, 2, \dots, n. \quad (115)$$

*Proof.* Suppose that  $X \leq_{\text{gdmftr}} Z$  w.r.t.  $\psi$ . This asserts that

$$\int_0^t \psi[F_X(x)] \left[ \frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} - \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \right] dx \geq 0, \quad \text{for all } t > 0. \quad (116)$$

By taking  $\psi(u) = B_{n-k+1,k}(u)$ ,  $u \in (0, 1)$ , we get that

$$\int_0^t \bar{B}_{n-k+1,k}(F_X(x)) \left[ \frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} - \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \right] dx \geq 0, \quad \text{for all } t > 0. \quad (117)$$

In view of the fact that

$$B_{n-k+1,k}(\bar{F}_X(x)) = \bar{F}_{X_{k:n}}(x),$$

$$\frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} = \frac{f_{X_{k:n}}(x)}{h_{Z_{k:n}}[H_{Z_{k:n}}^{-1}(F_{X_{k:n}}(x))]}, \quad (118)$$

thus

$$\int_0^t \bar{F}_{X_{k:n}}(x) \left[ \frac{f_{X_{k:n}}(t)}{h_{Z_{k:n}}[H_{Z_{k:n}}^{-1}(F_{X_{k:n}}(t))]} - \frac{f_{X_{k:n}}(x)}{h_{Z_{k:n}}[H_{Z_{k:n}}^{-1}(F_{X_{k:n}}(x))]} \right] dx \geq 0, \quad \text{for all } t > 0, \quad (119)$$

which is equivalent to

$$X_{k:n} \leq_{\text{dmtfr}} Z_{k:n}, \quad \text{for all } k = 1, 2, \dots, n. \quad (120)$$

Hence, the proof is complete.

Let  $k = 1$  and  $k = n$  in Theorem 12, respectively, we have the following corollary.  $\square$

**Corollary 1.** Assume that  $X \leq_{\text{gdmftr}} Z$  w.r.t.  $\psi$ .

- (i) If  $\psi(u) = B_{n,1}(u) = \int_0^u nx^{n-1} dx = u^n$ ,  $u \in (0, 1)$ . Then,  $X_{1:n} \leq_{\text{dmtfr}} Z_{1:n}$ .
- (ii) If  $\psi(u) = B_{1,n}(u) = \int_0^u n(1-x)^{n-1} dx = 1 - (1-u)^n$ ,  $u \in (0, 1)$ . Then  $X_{n:n} \leq_{\text{dmtfr}} Z_{n:n}$ .

**Theorem 13.** Assume that  $X_1, X_2, \dots$ , and  $Z_1, Z_2, \dots$ , are the i.i.d. copies of  $X$  and  $Z$ , respectively. Let  $N$  be a positive integer-valued random variable that is independent of  $X_i$ 's and  $Y_i$ 's.

(1) Let the function  $\psi$  be star-shaped with respect to the point  $(1, 0)$ .

- (i) If  $X_{1:N} \leq_{\text{dmtfr}} Z_{1:N}$ , then  $X_{1:N} \leq_{\text{gdmftr}} Z_{1:N}$  w.r.t.  $\psi$ .
- (ii) If  $X_{N:N} \leq_{\text{dmtfr}} Z_{N:N}$ , then  $X_{N:N} \leq_{\text{gdmftr}} Z_{N:N}$  w.r.t.  $\psi$ .

(2) Let the function  $\psi$  be anti-star-shaped with respect to the point  $(1, 0)$ .

- (i) If  $X_{1:N} \leq_{\text{gdmftr}} Z_{1:N}$  w.r.t.  $\psi$ , then  $X_{1:N} \leq_{\text{dmtfr}} Z_{1:N}$ .
- (ii) If  $X_{N:N} \leq_{\text{gdmftr}} Z_{N:N}$  w.r.t.  $\psi$ , then  $X_{N:N} \leq_{\text{dmtfr}} Z_{N:N}$ .

*Proof.* We only give the proof for the case (1) (b), the proofs for other cases are similar and hence omitted. If  $X_{N:N} \leq_{\text{dmtfr}} Z_{N:N}$ , from (30), we obtain that, for all  $t \geq 0$ ,

$$\int_0^t \bar{F}_{X_{N:N}}(x) \cdot \left[ \frac{f_{X_{N:N}}(t)}{h_{Z_{N:N}}[H_{Z_{N:N}}^{-1}(F_{X_{N:N}}(t))]} - \frac{f_{X_{N:N}}(x)}{h_{Z_{N:N}}[H_{Z_{N:N}}^{-1}(F_{X_{N:N}}(x))]} \right] dx \geq 0. \quad (121)$$

Since the function  $\psi$  is star-shaped with respect to the point  $(1, 0)$ , then the function

$$\frac{\psi(u)}{(1-u)}, \quad (122)$$

is decreasing in  $u \in (0, 1)$ . Hence,  $\psi[F_{X_{N:n}}(x)]/\bar{F}_{X_{N:n}}(x)$  is decreasing in  $x \geq 0$ ; from Lemma 4 and inequality (121), we have for all  $t \geq 0$ ,

$$\int_t^\infty \psi[F_{X_{N:n}}(x)] \cdot \left[ \frac{f_{X_{N:n}}(t)}{h_{Z_{N:n}}[H_{Z_{N:n}}^{-1}(F_{X_{N:n}}(t))]} - \frac{f_{X_{N:n}}(x)}{h_{Z_{N:n}}[H_{Z_{N:n}}^{-1}(F_{X_{N:n}}(x))]} \right] dx \geq 0. \quad (123)$$

That is,  $X_{N:n} \leq_{\text{gdmftr}} Z_{N:n}$  w.r.t.  $\psi$  as claimed.  $\square$

(2) Let the function  $\psi$  be dual anti-star-shaped of order  $n$ . If  $X_{n:n} \leq_{\text{dmftr}} Z_{n:n}$ , then  $X \leq_{\text{gdmftr}} Z$  w.r.t.  $\psi$ .

#### Theorem 14

(1) Let the function  $\psi$  be star-shaped of order  $n$  with respect to the point  $(1, 0)$ . If  $X_{1:n} \leq_{\text{dmftr}} Z_{1:n}$ , then  $X \leq_{\text{gdmftr}} Z$  w.r.t.  $\psi$ .

*Proof.* We only give the proof for case (1), the proof for case (2) is similar, and hence omitted. If  $X_{1:n} \leq_{\text{dmftr}} Z_{1:n}$ , from (30), we obtain that

$$\int_0^t \bar{F}_{X_{1:n}}(x) \cdot \left[ \frac{f_{X_{1:n}}(t)}{h_{Z_{1:n}}[H_{Z_{1:n}}^{-1}(F_{X_{1:n}}(t))]} - \frac{f_{X_{1:n}}(x)}{h_{Z_{1:n}}[H_{Z_{1:n}}^{-1}(F_{X_{1:n}}(x))]} \right] dx \geq 0, \quad \text{for all } t \geq 0. \quad (124)$$

Shaked and Shanthikumar [8] showed that  $H_{Z_{n:n}}^{-1}(F_{X_{n:n}}(x)) = H_Z^{-1}(F_X(x))$ . It can be proven that

$$\frac{f_{X_{1:n}}(x)}{h_{Z_{1:n}}[H_{Z_{1:n}}^{-1}(F_{X_{1:n}}(x))]} = \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]}, \quad \text{for all } x \geq 0. \quad (125)$$

Since the function  $\psi(x)$  is star-shaped of order  $n$  with respect to the point  $(1, 0)$ , then the function  $\psi(x)/(1-x)^n$  is decreasing in  $x \in (0, 1)$ . Hence, the function

$$\frac{\psi[F_X(x)]}{[\bar{F}_X(x)]^n}, \quad (126)$$

is nonnegatively decreasing in  $x \geq 0$ . From Lemma 4, inequality (124), and equation (125), we have

$$\int_0^t \psi[F_X(x)] \cdot \left[ \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} - \frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} \right] dx \geq 0, \quad \text{for all } t > 0. \quad (127)$$

That is,  $X \leq_{\text{gdmftr}} Z$  w.r.t.  $\psi$  as claimed.

At the end of this section, we show the relation between the orders DMFTR, RDMFTR, and GDMFTR:

$$\begin{aligned} X \leq_{\text{gdmftr}} Z \text{ w.r.t. } \psi, \quad \psi(u) = 1-u, u \in (0, 1) &\iff X \leq_{\text{dmftr}} Z, \\ X \leq_{\text{gdmftr}} Z \text{ w.r.t. } \psi, \quad \psi(u) = g_Y[G_Y^{-1}(u)], u \in (0, 1) &\iff X \leq_{\text{rdmftr}} Z, \\ X \leq_{\text{rdmftr}} Z \text{ w.r.t. } Y, \quad Y \text{ is exponential} &\iff X \leq_{\text{dmftr}} Z. \end{aligned} \quad (128)$$

## 4. Closure and Reversed Closure Properties of the GDMFTR Order

In this section, we investigate the closure and reversed closure properties of the GDMFTR order with respect to the

referred functions and the compared random variables, respectively.

The following Theorem 15 indicates that the GDMFTR order has the closure property with respect to the referred function under multiplication of functions.

**Theorem 15.** If  $X \leq_{\text{gdmftr}} Z$  w.r.t.  $\psi$ , and  $\varphi$  is nonnegative decreasing in  $x \in (0, 1)$ , then

$$X \leq_{\text{gdmftr}} Z \text{ w.r.t. } \varphi \cdot \psi. \quad (129)$$

*Proof.* Suppose that  $X \leq_{\text{gdmftr}} Z$  w.r.t.  $\psi$ . Then, from (30), we have

$$\int_0^t \psi[F_X(x)] \left[ \frac{f_X(t)}{g_Y[G_Y^{-1}(F_X(t))]} - \frac{f_X(x)}{g_Y[G_Y^{-1}(F_X(x))]} \right] dx \geq 0, \quad \text{for all } t > 0. \quad (130)$$

Since  $\varphi[F_X(t)]$  is nonnegatively decreasing in  $x \in (0, 1)$ , from (130) and Lemma 4, we obtain for all  $t > 0$ ,

$$\int_0^t \varphi[F_X(x)] \cdot \psi[F_X(x)] \left[ \frac{f_X(t)}{g_Y[G_Y^{-1}(F_X(t))]} - \frac{f_X(x)}{g_Y[G_Y^{-1}(F_X(x))]} \right] dx \geq 0, \quad (131)$$

which is equivalent to that  $X \leq_{\text{gdmftr}} Z$  w.r.t.  $\psi \cdot \varphi$ . Therefore, the proof of the theorem is complete.

From Theorem 15, we have the following corollary.  $\square$

**Corollary 2.** If  $X \leq_{\text{gdmftr}} Z$  w.r.t.  $\psi$ , and  $\psi$  is decreasing in  $x \in (0, 1)$ , then

$$X \leq_{\text{gdmftr}} Z \text{ w.r.t. } \psi^n. \quad (132)$$

**Theorem 16.** Let  $\psi_1$  and  $\psi_2$  be two continuous nonnegative functions defined on  $(0, 1)$ ,  $X$  and  $Z$  be two continuous nonnegative random variables. Then, the following statements are true.

(1) If  $X \leq_{\text{gdmftr}} Z$  w.r.t.  $\psi_i$ ,  $i = 1, 2$ , then

$$X \leq_{\text{gdmftr}} Z \text{ w.r.t. } \psi_1 + \psi_2. \quad (133)$$

(2) If  $X \leq_{\text{gdmftr}} Z$  w.r.t.  $\psi_1$  and  $\psi_2$  is increasing in  $x \in (0, 1)$ , then

$$X \leq_{\text{gdmftr}} Z \text{ w.r.t. } \frac{\psi_1}{\psi_2}. \quad (134)$$

(3) If  $X \leq_{\text{gdmftr}} Z$  w.r.t.  $\psi_2$ ,  $\psi_1$  is star-shaped and  $\psi_2$  is decreasing, then

$$X \leq_{\text{gdmftr}} Z \text{ w.r.t. } \psi_1 \circ \psi_2, \quad (135)$$

where  $\psi_1 \circ \psi_2$  is the compound of  $\psi_1$  and  $\psi_2$ .

(4) If  $X \leq_{\text{gdmftr}} Z$  w.r.t.  $\psi_2$ ,  $\psi_1$  is anti-star-shaped and  $\psi_2$  is increasing, then

$$X \leq_{\text{gdmftr}} Z \text{ w.r.t. } \psi_1 \circ \psi_2, \quad (136)$$

(5) If  $X \leq_{\text{gdmftr}} Z$  w.r.t.  $\psi_1/\psi_2$  and  $\psi_2$  is decreasing in  $x \in (0, 1)$ , then

$$X \leq_{\text{gdmftr}} Z \text{ w.r.t. } \psi_1. \quad (137)$$

(6) If  $X \leq_{\text{gdmftr}} Z$  w.r.t.  $\psi_2$  and  $\psi_1/\psi_2$  is decreasing in  $x \in (0, 1)$ , then

$$X \leq_{\text{gdmftr}} Z \text{ w.r.t. } \psi_1. \quad (138)$$

*Proof.* We only give the proofs for the case of (3); the proofs for the other cases are similar and omitted. (3) Suppose that  $X \leq_{\text{gdmftr}} Z$  w.r.t.  $\psi_2$ . Then, from (30), we have

$$\int_0^t \psi_2[F_X(x)] \left[ \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} - \frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} \right] dx \geq 0, \quad \text{for all } t > 0. \quad (139)$$

Since the function  $\psi_2$  is decreasing,  $\psi_2[F_X(x)]$  is then decreasing. Moreover,  $\psi_1$  is star-shaped on the interval  $(0, 1)$ , and then  $\psi_1(u)/u$  is increasing in  $u \in (0, 1)$ ; thus,

the function  $\psi_1[\psi_2(F_X(x))]/\psi_2[F_X(x)]$  is decreasing in  $x \geq 0$ . From Lemma 4 and inequality (139), we have for all  $t > 0$ ,



$$\int_0^t \psi_1[\psi_2(F_X(x))] \left[ \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} - \frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} \right] dx \geq 0, \quad (140)$$

which asserts that  $X \leq_{\text{gdmtr}} Z$  r.t.  $\psi_1 \circ \psi_2$  as claimed.

Let  $X$  be a nonnegative absolutely continuous random variable and  $\phi$  be a nonnegative increasing function defined on  $[0, \infty)$  with  $\phi(0) = 0$ . We call that  $\phi$  is a generalized scale function and  $\phi(X)$  is the generalized scale transform of  $X$ . In the accelerated life testing, the  $\phi$  is called an accelerating factor.

If a function  $\phi$  is increasing convex with  $\phi(0) = 0$ , then  $\phi$  is called a risk preference function, and  $\phi(X)$  is called the risk preference transform of  $X$ . If  $\phi$  is increasing concave with  $\phi(0) = 0$ , then  $\phi$  is called a risk aversion function, and  $\phi(X)$  is called the risk aversion transform of  $X$ . If  $\phi$  is a risk aversion function, we say that  $-\phi''/\phi'$  is the absolute risk aversion coefficient of  $\phi$ . If  $\phi$  is a risk preference function, we say that  $-\phi''/\phi'$  is the risk preference coefficient of  $\phi$ .  $\square$

**Theorem 17.** Let  $\psi$  be a nonnegative function defined on  $(0, 1)$  and  $\phi$  be a nonnegative function defined on  $[0, +\infty)$

with  $\phi(0) = 0$ . Suppose that the function  $\phi$  is increasing concave with  $-\phi''/\phi'$  being decreasing in  $x \geq 0$ , and  $X \geq_{\text{disp}} Z$ . Then,

$$X \leq_{\text{gdmtr}} Z \text{ w.r.t. } \psi \Rightarrow \phi(X) \leq_{\text{gdmtr}} \phi(Z) \text{ w.r.t. } \psi. \quad (141)$$

*Proof.* Suppose that  $X \leq_{\text{gdmtr}} Z$  w.r.t.  $\psi$ . Then, from (30), we have for all  $t > 0$ ,

$$\int_0^t \psi[F_X(x)] \left[ \frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} - \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \right] dx \geq 0. \quad (142)$$

Since the function  $\phi(x)$  is increasing concave,  $\phi'(x)$  is nonnegatively decreasing in  $x \geq 0$ . From Lemma 4 and (142), we have

$$\int_0^t \psi[F_X(x)] \phi'(x) \left[ \frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} - \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \right] dx \geq 0, \quad \text{for all } t > 0. \quad (143)$$

On the other hand,  $\phi(X) \leq_{\text{gdmtr}} \phi(Z)$  w.r.t.  $\psi$  if and only if for all  $t > 0$ ,

$$\begin{aligned} & \int_0^t \psi[F_{\phi(X)}(x)] \left[ \frac{f_{\phi(X)}(t)}{h_{\phi(Z)}[H_{\phi(Z)}^{-1}(F_{\phi(X)}(t))]} - \frac{f_{\phi(X)}(x)}{h_{\phi(Z)}[H_{\phi(Z)}^{-1}(F_{\phi(X)}(x))]} \right] \\ &= \int_0^t \psi[F_X(x)] \phi'(x) \left[ \frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} \cdot \frac{\phi'[\alpha(t)]}{\phi'(t)} - \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \cdot \frac{\phi'[\alpha(x)]}{\phi'(x)} \right] dx \geq 0, \end{aligned} \quad (144)$$

where  $\alpha(x) = H_Z^{-1}(F_X(x))$ . Under the given conditions of the theorem, we have that the function

$$\frac{\phi'[\alpha(x)]}{\phi'(x)}, \quad (145)$$

is increasing and nonnegative for all  $x > 0$ . Hence, from (143), we obtain for all  $t > 0$ ,

$$\begin{aligned} & \int_0^t \psi[F_X(x)] \phi'(x) \left[ \frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} \cdot \frac{\phi'[\alpha(t)]}{\phi'(t)} - \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \cdot \frac{\phi'[\alpha(x)]}{\phi'(x)} \right] dx \geq \\ & \frac{\phi'[\alpha(t)]}{\phi'(t)} \int_0^t \psi[F_X(x)] \phi'(x) \left[ \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} - \frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} \right] dx \geq 0. \end{aligned} \quad (146)$$

That is to say, inequality (144) is valid, so the proof of the theorem is complete.  $\square$

*Remark 7.* Theorem 17 says, under some appropriate conditions, that the order  $\leq_{\text{gdmtr}}$  has the closure property under the concave generalized scale transform.

*Remark 8.* Theorem 17 also says, under some appropriate conditions, that the order  $\leq_{\text{gdmtr}}$  has the closure property under the action of an increasing concave accelerating factor.

*Remark 9.* Theorem 17 also says, under the condition of  $X \geq_{\text{disp}} Z$ , that the GDMTFR order has the closure property under the risk averse transform  $\phi$  with an increasing absolute risk aversion coefficient  $-\phi''/\phi'$ .

With a similar manner of the proof of Theorem 17, Theorem 18 can be proven, and the detailed proof is omitted.

**Theorem 18.** Let  $\psi$  be a nonnegative function defined on  $(0, 1)$  and  $\phi$  be a nonnegative function defined on  $[0, +\infty)$  with  $\phi(0) = 0$ . Suppose that the function  $\phi$  is increasing and convex with  $\phi''/\phi'$  being increasing in  $x \geq 0$ . If  $\phi(X) \leq_{\text{gdmtr}} \phi(Z)$  w.r.t.  $\psi$  and  $X \geq_{\text{disp}} Z$ , then

$$X \leq_{\text{gdmtr}} Z \text{ w.r.t. } \psi. \quad (147)$$

*Remark 10.* Theorem 18 says, under some appropriate conditions, that the order  $\leq_{\text{gdmtr}}$  has the reversed closure property under the convex generalized scale transform.

*Remark 11.* Theorem 18 also says, under some appropriate conditions, that the order  $\leq_{\text{gdmtr}}$  has the reversed closure property under the action of an increasing and convex accelerating factor.

*Remark 12.* Theorem 18 also says, under the condition of  $X \geq_{\text{disp}} Z$ , that the GDMTFR order has the reversed closure property under the risk preference transform  $\phi$  with an increasing risk preference coefficient  $\phi''/\phi'$ .

**Theorem 19.** Let  $\psi$  be a nonnegative increasing function defined on  $(0, 1)$  and let  $\phi$  be a nonnegative increasing concave function defined on  $[0, +\infty)$  with  $\phi(0) = 0$  and with  $\phi''/\phi'$  being decreasing in  $x > 0$ . Suppose that  $X \geq_{\text{disp}} Z$ . Then,

$$X \leq_{\text{gdmtr}} Z \text{ w.r.t. } \psi \Rightarrow \phi(X) \leq_{\text{gdmtr}} \phi(Z) \text{ w.r.t. } \phi \circ \psi. \quad (148)$$

*Proof.* Suppose that  $X \leq_{\text{gdmtr}} Z$  w.r.t.  $\psi$ . Then, from (30), we have

$$\int_0^t \psi[F_X(x)] \left[ \frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} - \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \right] dx \geq 0, \quad \text{for all } t > 0. \quad (149)$$

Since the function  $\phi(x)$  is increasing concave,  $\phi'(x)$  is nonnegatively decreasing in  $x > 0$ . From Lemma 4 and (149), we have

$$\int_0^t \psi[F_X(x)] \phi'(x) \left[ \frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} - \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \right] dx \geq 0, \quad \text{for all } t > 0. \quad (150)$$

Since  $\phi$  is nonnegatively increasing concave,  $\phi(x)/x$  is nonnegatively decreasing and  $\psi$  being nonnegatively increasing leads to increasing  $\psi(F_X(x))$ , so

$\phi[\psi(F_X(x))]/\psi(F_X(x))$  is nonnegatively decreasing. From Lemma 4 and (150), we obtain that

$$\int_0^t \phi[\psi(F_X(x))] \phi'(x) \left[ \frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} - \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \right] dx \geq 0, \quad \text{for all } t > 0. \quad (151)$$

On the other hand,  $\phi(X) \leq_{\text{gdmtr}} \phi(Z)$  w.r.t.  $\phi \circ \psi$  if and only if for all  $t \geq 0$ ,

$$\begin{aligned} & \int_0^t \phi[\psi(F_{\phi(X)}(x))] \left[ \frac{f_{\phi(X)}(t)}{h_{\phi(Z)}[H_{\phi(Z)}^{-1}(F_{\phi(X)}(t))]} - \frac{f_{\phi(X)}(x)}{h_{\phi(Z)}[H_{\phi(Z)}^{-1}(F_{\phi(X)}(x))]} \right] \\ &= \int_0^t \phi[\psi(F_X(x))] \phi'(x) \left[ \frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} \cdot \frac{\phi'[\alpha(t)]}{\phi'(t)} - \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \cdot \frac{\phi'[\alpha(x)]}{\phi'(x)} \right] dx \geq 0, \end{aligned} \quad (152)$$

where  $\alpha(x) = H_Z^{-1}(F_X(x))$ . Since  $\phi''/\phi'$  is decreasing, and  $X \geq_{\text{disp}} Z$ , we have that the function

$$\frac{\phi'[\alpha(x)]}{\phi'(x)}, \quad (153)$$

is nonnegatively increasing in  $x > 0$ . Therefore, for all  $t > 0$ ,

$$\begin{aligned} & \int_0^t \phi[\psi(F_X(x))] \phi'(x) \left[ \frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} \cdot \frac{\phi'[\alpha(t)]}{\phi'(t)} - \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \cdot \frac{\phi'[\alpha(x)]}{\phi'(x)} \right] dx \\ & \geq \frac{\phi'[\alpha(t)]}{\phi'(t)} \int_0^t \phi[\psi(F_X(x))] \phi'(x) \left[ \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} - \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(t))]} \right] dx \geq 0. \end{aligned} \quad (154)$$

From (151) and (154), we see that inequality (152) is valid. That is to say,  $\phi(X) \leq_{\text{gdmtr}} \phi(Z)$  w.r.t.  $\phi \circ \psi$ . Therefore, the proof of the theorem is complete.  $\square$

*Remark 13.* Theorem 19 says that, under some appropriate conditions, the order  $\leq_{\text{gdmtr}}$  has the closure property with respect to both the compared random variables and the referred function under the concave generalized scale transforms.

*Remark 14.* Theorem 19 also says that, under some appropriate conditions, the order  $\leq_{\text{gdmtr}}$  has the closure property with respect to both the compared random variables and the referred function under the action of a concave accelerating factor.

*Remark 15.* Theorem 19 also says that, under the condition of  $X \geq_{\text{disp}} Z$ , the GDMTR order has the closure property with respect to both the compared random variables and the referred function under the risk aversion transform  $\phi$  with an increasing absolute risk aversion coefficient  $-\phi''/\phi'$ .

**Theorem 20.** Let  $\psi$  be a nonnegative increasing function defined on  $(0, 1)$ , and let  $\phi$  be a nonnegative increasing convex function defined on  $[0, +\infty)$  with  $\phi(0) = 0$  and with  $-\phi''/\phi'$  being decreasing in  $x > 0$ . Assume that  $X \geq_{\text{disp}} Z$ . Then,

$$\phi(X) \leq_{\text{gdmtr}} \phi(Z) \text{ w.r.t. } \phi \circ \psi \Rightarrow X \leq_{\text{gdmtr}} Z \text{ w.r.t. } \psi. \quad (155)$$

*Proof.* Suppose that  $\phi(X) \leq_{\text{gdmtr}} \phi(Z)$  w.r.t.  $\phi \circ \psi$ . Then, from (30), we have, for all  $t > 0$ ,

$$\begin{aligned} & \int_0^t \phi[\psi(F_{\phi(X)}(x))] \left[ \frac{f_{\phi(X)}(t)}{h_{\phi(Z)}[H_{\phi(Z)}^{-1}(F_{\phi(X)}(t))]} - \frac{f_{\phi(X)}(x)}{h_{\phi(Z)}[H_{\phi(Z)}^{-1}(F_{\phi(X)}(x))]} \right] \\ &= \int_0^t \phi[\psi(F_X(x))] \phi'(x) \left[ \frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} \cdot \frac{\phi'[\alpha(t)]}{\phi'(t)} - \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \cdot \frac{\phi'[\alpha(x)]}{\phi'(x)} \right] dx \geq 0. \end{aligned} \quad (156)$$

Since  $\phi$  is increasing convex,  $\phi(y)/y$  is increasing, and  $\psi$  is increasing. This leads to  $\psi(F_X(x))$  is nonnegatively increasing. Combining these facts, we obtain that the function

$$\frac{\psi(F_X(x))}{\phi[\psi(F_X(x))]}, \quad (157)$$

is nonnegative decreasing in  $x > 0$ . Moreover, since  $\phi$  is increasing convex,  $1/\phi'$  is nonnegatively decreasing. So,

$$\int_0^t \psi(F_X(x)) \left[ \frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} \cdot \frac{\phi'[\alpha(t)]}{\phi'(t)} - \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \cdot \frac{\phi'[\alpha(x)]}{\phi'(x)} \right] dx \geq 0. \quad (159)$$

Under the given conditions that  $\phi$  is increasing convex with  $\phi''/\phi'$  being decreasing in  $x \geq 0$ , and  $X \geq_{\text{disp}} Z$ , we get

$$\begin{aligned} 0 &\leq \int_0^t \psi(F_X(x)) \left[ \frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} \cdot \frac{\phi'[\alpha(t)]}{\phi'(t)} - \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \cdot \frac{\phi'[\alpha(x)]}{\phi'(x)} \right] dx \\ &\leq \frac{\phi'[\alpha(t)]}{\phi'(t)} \int_0^t \psi(F_X(x)) \left[ \frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} - \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \right] dx, \quad \text{for all } t > 0. \end{aligned} \quad (160)$$

Thus,

$$\int_0^t \psi[F_X(x)] \left[ \frac{f_X(t)}{h_Z[H_Z^{-1}(F_X(t))]} - \frac{f_X(x)}{h_Z[H_Z^{-1}(F_X(x))]} \right] dx \geq 0. \quad (161)$$

That is to say,  $\phi(X) \leq_{\text{gdmtr}} \phi(Z)$  w.r.t.  $\phi \circ \psi$ . Therefore, the proof of the theorem is complete.  $\square$

*Remark 16.* Theorem 20 says that, under some appropriate conditions, the order  $\leq_{\text{gdmtr}}$  has the reversed closure property with respect to both the compared random variables and the referred function under the convex generalized scale transforms.

*Remark 17.* Theorem 20 also says that, under some appropriate conditions, the order  $\leq_{\text{gdmtr}}$  has the reversed closure property with respect to both the compared random variables and the referred function under the action of a convex accelerating factor.

*Remark 18.* Theorem 20 also says that, under the condition of  $X \geq_{\text{disp}} Z$ , the GDMTFR order has the reversed closure property with respect to both the compared random variables and the referred function under the risk preference transform  $\phi$  with a decreasing risk preference coefficient  $\phi''/\phi'$ .

$$\frac{\psi(F_X(x))}{\phi[\psi(F_X(x))]\phi'(x)}, \quad (158)$$

is nonnegatively decreasing in  $x > 0$ . From Lemma 4 and (156), we obtain that, for all  $t > 0$ ,

that  $\phi'[\alpha(x)]/\phi'(x)$  is nonnegatively decreasing in  $x > 0$ . Hence,

## 5. Examples

We give several illustrative examples that meet the GDMTFR order.

*Example 1.* Let  $\psi$  be a nonnegative function defined on the interval  $(0, 1)$ . Now, we consider comparing two Weibull random variables in the GDMTFR order. Let  $X_i, i = 1, 2$ , be two Weibull random variables with respective survival function  $\bar{F}_{X_i}$ . Specifically, let  $X_i \sim W(\alpha_i, \lambda)$ ,  $i = 1, 2$ ,  $\bar{F}_{X_i}(x) = e^{-\lambda x^{\alpha_i}}$ ,  $x \geq 0$ . Hence, the density functions of  $X_i$  are, respectively,

$$f_{X_i}(x) = \lambda \alpha_i x^{\alpha_i - 1} e^{-\lambda x^{\alpha_i}}, \quad x \geq 0. \quad (162)$$

Then, the GDMTFR order is determined by the parameters  $\alpha_i$ . Indeed, it is easy to verify that

$$\begin{aligned} F_{X_2}^{-1}[F_{X_1}(x)] &= x^{\alpha_1/\alpha_2}, \\ f_{X_2}(x) &= \lambda \alpha_2 x^{\alpha_2 - 1} e^{-\lambda x^{\alpha_2}}. \end{aligned} \quad (163)$$

Hence,

$$\frac{f_{X_1}(x)}{f_{X_2}[F_{X_2}^{-1}(F_{X_1}(x))]} = \frac{\alpha_1}{\alpha_2} x^{\alpha_1/\alpha_2 - 1}. \quad (164)$$

(1) If  $\alpha_1 \geq \alpha_2 > 0$ , then from (164), it is clear that the function  $f_{X_1}(x)/f_{X_2}[F_{X_2}^{-1}(F_{X_1}(x))]$  is increasing in

$x > 0$ . From this fact and (30), we have that  $X_1 \leq_{\text{gdmftr}} X_2$  w.r.t.  $\psi$ .

- (2) If  $0 < \alpha_1 \leq \alpha_2$ , then from (164), it is clear that the function  $f_{X_1}(x)/f_{X_2}[F_{X_2}^{-1}(F_{X_1}(x))]$  is decreasing in  $x > 0$ . From this fact and (21), we have that  $X_1 \geq_{\text{gdmftr}} X_2$  w.r.t.  $\psi$ .

*Example 2.* Consider a comparison of two beta random variables in the GDMTFR order. Let  $\psi$  be a nonnegative function defined on the interval  $(0, 1)$ . Let  $X_i, i = 1, 2$ , be two Beta random variables with respective distribution functions  $F_{X_i}$ . Specifically, let  $X_i \sim \text{Beta}(\alpha_i, 1)$ ,  $F_{X_i}(x) = x^{\alpha_i}$ ,  $x \in (0, 1)$ . Hence, the density functions of  $X_i$  are given by, respectively,

$$f_{X_i}(x) = \alpha_i x^{\alpha_i - 1}, \quad x \in (0, 1). \quad (165)$$

It is readily verified that

$$\begin{aligned} F_{X_2}^{-1}[F_{X_1}(x)] &= x^{\alpha_1/\alpha_2}, \\ f_{X_2}(x) &= \lambda \alpha_2 x^{\alpha_2 - 1} e^{-\lambda x^{\alpha_2}}, \quad x \in (0, 1). \end{aligned} \quad (166)$$

Hence,

$$\frac{f_{X_1}(x)}{f_{X_2}[F_{X_2}^{-1}(F_{X_1}(x))]} = \frac{\alpha_1}{\alpha_2} x^{(\alpha_1/\alpha_2) - 1}. \quad (167)$$

- (1) If  $\alpha_1 \geq \alpha_2 > 0$ , it can be seen that the function  $f_{X_1}(x)/f_{X_2}[F_{X_2}^{-1}(F_{X_1}(x))]$  is increasing in  $x \geq 0$ . From this fact and (30), we see that  $X_1 \leq_{\text{gdmftr}} X_2$  w.r.t.  $\psi$ .
- (2) If  $0 < \alpha_1 \leq \alpha_2$ , from (171), we see that the function  $f_{X_1}(x)/f_{X_2}[F_{X_2}^{-1}(F_{X_1}(x))]$  is decreasing in  $x \geq 0$ . From this fact and (30), we have that  $X_1 \geq_{\text{gdmftr}} X_2$  w.r.t.  $\psi$ .

*Example 3.* Consider a comparison of another two beta random variables in the GDMTFR order. Let  $\psi$  be a nonnegative function defined on the interval  $(0, 1)$ . Let  $X_i, i = 1, 2$ , be two beta random variables with respective distribution function  $F_{X_i}$ . Specifically, let  $X_i \sim \text{Beta}(1, \beta_i)$ , then the density functions of  $X_i$  are given by, respectively,

$$f_{X_i}(x) = \beta_i (1-x)^{\beta_i - 1}, \quad x \in (0, 1), \quad i = 1, 2. \quad (168)$$

The distribution functions of  $X_i$  are thus

$$F_{X_i}(x) = 1 - (1-x)^{\beta_i}, \quad x \in (0, 1), \quad i = 1, 2, \quad (169)$$

respectively.

It can be verified that

$$\begin{aligned} F_{X_2}^{-1}[F_{X_1}(x)] &= 1 - (1-x)^{\beta_1/\beta_2}, \\ f_{X_2}(x) &= \beta_2 (1-x)^{\beta_2 - 1}, \quad x \in (0, 1). \end{aligned} \quad (170)$$

Hence,

$$\frac{f_{X_1}(x)}{f_{X_2}[F_{X_2}^{-1}(F_{X_1}(x))]} = \frac{\beta_1}{\beta_2} (1-x)^{(\beta_1/\beta_2) - 1}. \quad (171)$$

- (1) If  $\beta_1 \geq \beta_2 > 0$ , it is clear that the function  $f_{X_1}(x)/f_{X_2}[F_{X_2}^{-1}(F_{X_1}(x))]$  is decreasing in  $x > 0$ . From this fact and (30), we see that  $X_1 \geq_{\text{gdmftr}} X_2$  w.r.t.  $\psi$ .
- (2) Assume that  $0 < \beta_1 \leq \beta_2$ . From (171), we see that  $f_{X_1}(x)/f_{X_2}[F_{X_2}^{-1}(F_{X_1}(x))]$  is increasing in  $x > 0$ . From this fact and (30), we see that  $X_1 \leq_{\text{gdmftr}} X_2$  w.r.t.  $\psi$ .

*Example 4.* Let  $\psi$  be a nonnegative function defined on the interval  $(0, 1)$ ; we consider to compare two Pareto random variables in the GDMTFR order. Specifically, assume that

$$X \sim \mathbb{P}(\alpha_1, \lambda), \quad Y \sim \mathbb{P}(\alpha_2, \lambda), \quad (172)$$

where  $\alpha_1, \alpha_2, \lambda$  are positive real numbers. That is, for all  $x \geq 0$ ,

$$\bar{F}_X(x) = \left( \frac{\lambda}{\lambda + x} \right)^{\alpha_1}, \quad (173)$$

$$\bar{G}_Y(x) = \left( \frac{\lambda}{\lambda + x} \right)^{\alpha_2},$$

$$f_X(x) = \frac{\alpha_1}{\lambda} \left( \frac{\lambda}{\lambda + x} \right)^{\alpha_1 + 1}, \quad (174)$$

$$g_Y(x) = \frac{\alpha_2}{\lambda} \left( \frac{\lambda}{\lambda + x} \right)^{\alpha_2 + 1}.$$

Then,

$$G_Y^{-1}(F_X(x)) = \bar{G}_Y^{-1}(\bar{F}_X(x)) = \lambda \left( 1 + \frac{x}{\lambda} \right)^{\alpha_1/\alpha_2} - \lambda. \quad (175)$$

Hence,

$$\frac{f_X(x)}{g_Y[G_Y^{-1}(F_X(x))]} = \frac{\alpha_1}{\alpha_2} \left( 1 + \frac{x}{\lambda} \right)^{(\alpha_1/\alpha_2) - 1}. \quad (176)$$

- (1) If  $\alpha_1 \geq \alpha_2 > 0$ , then from (176), we see that the function  $f_X(x)/g_Y[G_Y^{-1}(F_X(x))]$  is increasing in  $x > 0$ ; hence, from (30), we have  $X \leq_{\text{gdmftr}} Y$  w.r.t.  $\psi$ .
- (2) If  $0 < \alpha_1 \leq \alpha_2$ , then  $f_X(x)/g_Y[G_Y^{-1}(F_X(x))]$  is decreasing in  $x > 0$ ; from (30), we get that  $X \geq_{\text{gdmftr}} Y$  w.r.t.  $\psi$ .

## Data Availability

In the examples of this research, we mainly use random variables to verify the correctness of the theory.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This paper was supported by the Fifth Round Project of Ningbo Social Science Research Base (grant no. JD5-ZD45), the scientific research and innovation team of Zhejiang Wanli University (grant no. 202036), and the Zhejiang Wanli University Teaching Reform Project “Applied University Financial Engineering Major: Approaches and Practices for First-class Curriculum Construction.”

## References

- [1] S. Ma, M. H. Islam, and S. Ahmed, “A case study on risk management in existing construction project in Bangladesh,” *Journal of Logistics, Informatics and Service Science*, vol. 5, no. 1, pp. 1–16, 2018.
- [2] H. Kabir, I. Hoque, and T. Rahman, “Engineering properties and cost comparison among sylhet sand, khustia sand and local sand in the context of foundation,” *Journal of Logistics, Informatics and Service Science*, vol. 5, no. 1, pp. 17–30, 2018.
- [3] J. Maria Fernandez-Ponce, S. C. Kochar, and J. Muñoz-Perez, “Partial orderings of distributions based on right-spread functions,” *Journal of Applied Probability*, vol. 35, no. 1, pp. 221–228, 1998.
- [4] F. Belzunce, “On a characterization of right spread order by the increasing convex order,” *Statistics & Probability Letters*, vol. 45, no. 2, pp. 103–110, 1999.
- [5] A. Müller and D. Stoyan, *Comparison Methods for Stochastic Models and Risks*, John Wiley & Sons Inc., New York, NY, USA, 2002.
- [6] X. Li and R. C. M. Yam, “Reversed preservation properties of some negative aging conceptions and stochastic orders,” *Statistical Papers*, vol. 46, no. 1, p. 65, 2005.
- [7] M. Shaked, J. G. Shanthikumar, and Y. L. Tong, “Stochastic orders and their applications,” *SIAM Review*, vol. 37, no. 3, pp. 477–478, 1995.
- [8] M. Shaked and J. G. Shanthikumar, *Stochastic Orders*, Springer Science & Business Media, Berlin, Germany, 2007.
- [9] P. Zhao and N. Balakrishnan, “Stochastic comparison and monotonicity of inactive record values,” *Statistics & Probability Letters*, vol. 79, no. 5, pp. 566–572, 2009.
- [10] S. M. Sunoj, P. G. Sankaran, and A. K. Nanda, “Quantile based entropy function in past lifetime,” *Statistics & Probability Letters*, vol. 83, no. 1, pp. 366–372, 2013.
- [11] D.-t. Kang, “Further results on closure properties of lpqe order,” *Statistical Methodology*, vol. 25, pp. 23–35, 2016.
- [12] D.-t. Kang, “Some new results on the lqe ordering,” *Statistical Methodology*, vol. 32, pp. 218–235, 2016.
- [13] L. Yan and D.-t. Kang, “Some new results on the rényi quantile entropy ordering,” *Statistical Methodology*, vol. 33, pp. 55–70, 2016.
- [14] D.-T. Kang and L. Yan, “On the dynamic cumulative residual quantile entropy ordering,” *Statistical Methodology*, vol. 32, no. 14–35, 2016.
- [15] B. Vineshkumar, N. U. Nair, and P. G. Sankaran, “Stochastic orders using quantile-based reliability functions,” *Journal of the Korean Surgical Society*, vol. 44, no. 2, pp. 221–231, 2015.
- [16] P. Kosec, S. Skec, and D. Miler, “A comparison of the tolerance analysis methods in the open-loop assembly,” *Advances in Production Engineering & Management*, vol. 15, no. 44–56, 2020.
- [17] H. Sun, Z. Lei, Y. Huang, and J. Yu, “Multi-objective robust design of new rotate barrel based on satisfaction function,” *Tehnički Vjesnik*, vol. 27, no. 2, pp. 558–566, 2020.
- [18] R. E. Barlow and F. Proschan, *Statistical Theory of Reliability and Life Testing*, Hold, Reinhart and Wiston, Inc., Silver Spring, MD, USA, 1981.
- [19] S. C. Kochar, X. Li, and M. Shaked, “The total time on test transform and the excess wealth stochastic orders of distributions,” *Advances in Applied Probability*, vol. 34, no. 4, pp. 826–845, 2002.
- [20] I. A. Ahmad, K. Mohamed, and P. Franco, “Further results involving the mit order and the imit class,” *Probability in the Engineering and Informational Sciences*, vol. 19, no. 3, pp. 377–395, 2005.
- [21] L. Knopik, “Some results on the ageing class,” *Control and Cybernetics*, vol. 34, pp. 1175–1180, 2005.
- [22] X. Li and M. Shaked, “A general family of univariate stochastic orders,” *Journal of Statistical Planning and Inference*, vol. 137, no. 11, pp. 3601–3610, 2007.
- [23] R. A. Barlow and R. E. Campo, “Total time on test processes and applications to failure data analysis,” *Reliability and Fault Tree Analysis*, University of California, Berkeley, CA, USA, 1975.
- [24] B. Klefsjö, “On aging properties and total time on test transforms,” *Scandinavian Journal of Statistics*, vol. 9, pp. 37–41, 1982.
- [25] A. W. Marshall, *Classes of Distributions Applicable in Replacement with Renewal Theory Implications*, University of California Press, Berkeley, CA, USA, 1972.
- [26] R. E. Barlow, “Geometry of the total time on test transform,” *Naval Research Logistics Quarterly*, vol. 26, no. 3, pp. 393–402, 1979.
- [27] A. W. Marshall and F. Proschan, “An inequality for convex functions involving majorization,” Technical Report, Boeing Scientific Research Labs, Seattle, WA, USA, 1964.
- [28] A. A. Emad-Eldin, “On testing for decreasing mean residual life ordering,” *Naval Research Logistics*, vol. 40, no. 5, pp. 633–642, 1993.
- [29] J. Bartoszewicz and M. Benduch, “Some properties of the generalized ttt transform,” *Journal of Statistical Planning and Inference*, vol. 139, no. 7, pp. 2208–2217, 2009.



## Retraction

# Retracted: Security and Privacy Risk Assessment of Energy Big Data in Cloud Environment

### Computational Intelligence and Neuroscience

Received 17 October 2023; Accepted 17 October 2023; Published 18 October 2023

Copyright © 2023 Computational Intelligence and Neuroscience. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] Z. Li, W. Xu, H. Shi, Y. Zhang, and Y. Yan, "Security and Privacy Risk Assessment of Energy Big Data in Cloud Environment," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 2398460, 11 pages, 2021.

## Research Article

# Security and Privacy Risk Assessment of Energy Big Data in Cloud Environment

Zhiru Li,<sup>1</sup> Wei Xu,<sup>1</sup> Huibin Shi,<sup>1</sup> Yuanyuan Zhang ,<sup>2</sup> and Yan Yan<sup>1</sup>

<sup>1</sup>Shenyang University of Technology, Shenyang, China

<sup>2</sup>Dalian Medical University, Dalian, China

Correspondence should be addressed to Yuanyuan Zhang; zhangyuan@dmu.edu.cn

Received 8 September 2021; Revised 21 September 2021; Accepted 27 September 2021; Published 14 October 2021

Academic Editor: Daqing Gong

Copyright © 2021 Zhiru Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Considering the importance of energy in our lives and its impact on other critical infrastructures, this paper starts from the whole life cycle of big data and divides the security and privacy risk factors of energy big data into five stages: data collection, data transmission, data storage, data use, and data destruction. Integrating into the consideration of cloud environment, this paper fully analyzes the risk factors of each stage and establishes a risk assessment index system for the security and privacy of energy big data. According to the different degrees of risk impact, AHP method is used to give indexes weights, genetic algorithm is used to optimize the initial weights and thresholds of BP neural network, and then the optimized weights and thresholds are given to BP neural network, and the evaluation samples in the database are used to train it. Then, the trained model is used to evaluate a case to verify the applicability of the model.

## 1. Introduction

In the era of big data, the application of big data technology in the energy field is a trend to promote industrial development and innovation. Both the deep application of big data technology in the energy field and the deep integration of energy production, consumption, and related technology revolution with big data concept will accelerate the development of energy industry [1].

With the implementation of the global energy big data strategy, the rapid development of “Internet plus” smart energy and the comprehensive construction of intelligent energy layout make the energy industry more widely distributed, more data collection points, more data types, more complex business relationships, and a wider range of data usage and users [2]. So while bringing convenience, it also brings risks to energy big data management. Due to the critical infrastructure of each country, energy is bound to become the preferred target of attack in case of cyber war. With the frequent occurrence of more and more energy security and privacy incidents, such as “Blackout in Ukraine” and “Stuxnet virus” attack on Iran’s nuclear facilities, big

data has become a usable and attachable carrier [3]. Through the big data value information obtained by the attack, the energy distribution of the target location can be analyzed, and the key data such as the monitoring and early warning information and operation instructions of key nodes will be tampered, resulting in energy system failure or major security accidents.

Therefore, the management research based on energy big data has been widely concerned by scholars all over the world. At present, for the huge amount of data and the particularity of management in the energy industry, scholars carry out data management and architecture design through various technical or nontechnical means, including the establishment of big data layer to store and process renewable energy data [4] and the establishment of energy big data processing system, supporting memory distributed computing [5]. In the research on the security and privacy of big data, it is found that most scholars used a single model for risk assessment, such as analytic hierarchy process (AHP), factor analysis, grey theory [6], fuzzy evaluation method [7], and cloud model [8]. Such methods are based on statistical theory and cannot completely get rid of the influence of

subjectivity and theoretical assumptions. In recent years, machine learning has become an important research tool in the field of security and privacy [9]. When using machine learning methods to evaluate and predict risks, the accuracy is often higher than that of traditional statistical methods [10]. Common machine learning methods include neural network, SVM, and clustering algorithm; BP neural network is the most widely used neural network in risk prediction and evaluation [11], but which is easy to fall into local minimum in practical application [12]. Therefore, scholars often use other algorithms as assistance to improve the accuracy of prediction and evaluation. For example, Zhang (2021) established a regression model through BP network and used PSO algorithm to optimize connection weights to evaluate the slow convergence of BP network, in order to improve the accuracy of rockburst prediction [13]. Wang (2019) et al. used LM algorithm to improve the operation efficiency and accuracy of traditional BP neural network and provided an effective theoretical basis and modeling method for risk prediction of power communication network [14].

This greatly improves the accuracy of prediction and evaluation, but a review of the relevant literature shows that the analysis of the importance of the impact of indexes is often neglected. Thus, in this paper, based on the consideration of machine learning, according to the different degrees of risk impact, AHP method is used to determine the index weight, which overcomes the deficiency of subjective consideration in previous studies [15]; the genetic algorithm optimized BP neural network (hereinafter referred to as GABP) with better prediction and evaluation effect is used for evaluation [16], which is a successful attempt to realize the combination of energy field and deep learning. In addition, for the security and privacy risk assessment of energy big data, the current literature pays more attention to theoretical analysis and lacks a relatively perfect assessment reference system. Starting from the whole life cycle of big data and considering the cloud environment, this paper establishes a risk assessment index system of energy big data security and privacy, which enriches the theoretical basis and framework in this field to a certain extent.

## 2. The Index System of Security and Privacy Risk Assessment of Energy Big Data in Cloud Environment

*2.1. Principles for the Construction of the Index System.* In the process of risk assessment, the probability of risk occurrence, loss range, and other factors need to be considered comprehensively to get the possibility and degree of system risk occurrence, determine the risk level, and then decide whether to take corresponding control measures and to what extent [17].

Therefore, the construction of risk assessment index system should follow the principles of comprehensiveness, scientificity, representativeness, and practicability, select the representative risk elements from a scientific perspective, quantify the risk based on the practical principle, and strive to show the risk management level comprehensively and accurately.

*2.2. Identification of Risk Factors.* Data security management is the most prominent risk faced by big data application. Although the massive data is stored centrally, it is convenient for data analysis and processing, but the loss and damage of big data caused by improper security management will cause devastating disaster. Due to the development of new technology and new business, the infringement of privacy right is not limited to physical and compulsory invasion, but is derived in a subtler way through various data, and the data security and privacy risks caused by this will be more serious [18].

Compared with the previous Internet and computer technology, the application advantage of big data in the cloud environment is more obvious. Big data platform has strong sharing ability, which can manage the security of information use and improve the efficiency of resource utilization. The construction of cloud platform and system application have strict standards. Cloud computing technology provides more comprehensive technical support and makes privacy management more reasonable, which is consistent with the level of technology development in the new era [19]. But from another point of view, it is under the influence of cloud platform sharing features that part of the data information is easy to be exposed, which provides opportunities for some illegal intrusion. Therefore, we must pay full attention to its risks.

Based on the literature of Xu [20], Tawalbeh [21], and He [22], combined with the analysis of relevant cases and the consultation of professionals, this paper follows the above evaluation index setting principle, combines with the development characteristics of energy big data security factors, and considers the impact of cloud environment. From the perspective of the whole life cycle of big data, this paper summarizes the current privacy security risks of cloud computing and big data and divides the risk assessment factors into five stages: data collection, data transmission, data storage, data use, and data destruction, with a total of 22 indexes, as shown in Figure 1.

*2.3. Index Quantification.* In terms of data collection, for the quantification of energy big data security and privacy risk indexes, this study introduces the concept of risk degree. According to the occurrence possibility and loss degree of each risk index, the product of possibility and loss degree is used as the reference standard of risk degree quantification, and the specific value can be reasonably floating around the product. The quantification of probability and loss degree can be divided into five levels: very high risk (5 points), high risk (4 points), medium risk (3 points), low risk (2 points), and very low risk (1 point).

$$R = P * L. \quad (1)$$

In formula (1),  $P$  is the probability of occurrence and  $L$  is the degree of loss.

The normalized input value is multiplied by the corresponding weight of each index as the input of the neural network for training, combined with the output value; the risk assessment level can be obtained, as shown in Table 1.

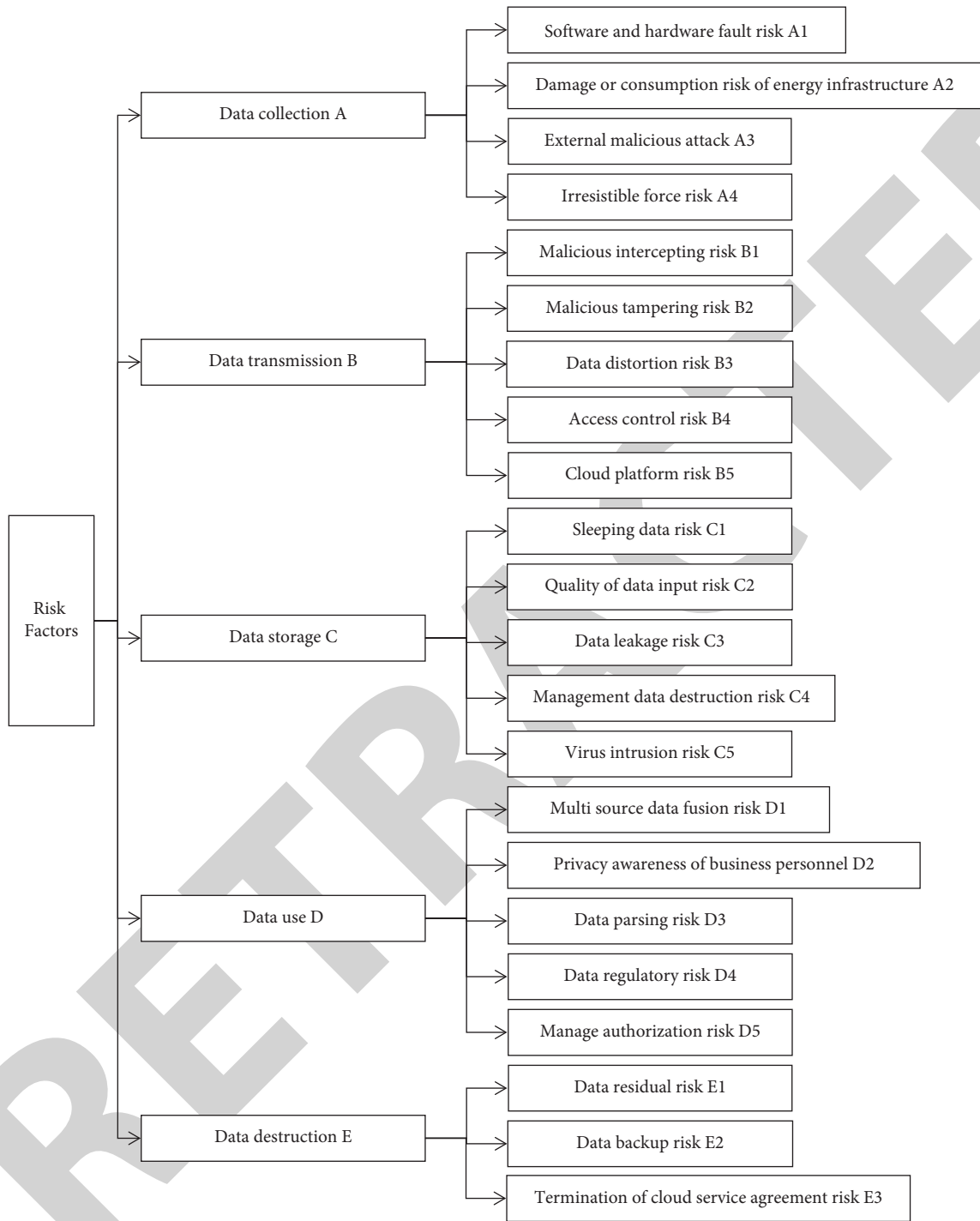


FIGURE 1: Security and privacy risk assessment index system based on the whole life cycle of energy big data.

TABLE 1: Risk assessment level.

Risk level	Meaning
First class ( $0 \leq R \leq 0.2$ )	The risk level is very low, so it is not necessary to pay special attention to it. The plan and general prevention can be made.
Second class ( $0.2 < R \leq 0.4$ )	The risk level is low, the plan and general prevention should be made, and need to be checked regularly.
Third class ( $0.4 < R \leq 0.6$ )	The risk level is medium; the major risk factors should be paid attention to in combination with the specific situation, and the corresponding countermeasures should be formulated.

TABLE 1: Continued.

Risk level	Meaning
Fourth class ( $0.6 < R \leq 0.8$ )	The risk level is high; it is necessary to pay attention to all the risk factors that may threaten the security of energy data, formulate the process sequence after the occurrence of the risk according to the importance degree, and track the inspection and evaluation.
Fifth class ( $0.8 < R \leq 1.0$ )	The risk level is very high; if necessary, it can be stopped and maintained, and the comprehensive inspection and special evaluation should be carried out immediately and can be continued after improvement.

### 3. Assessment Model of Energy Big Data Security and Privacy Risk in Cloud Environment

**3.1. AHP Method.** In the existing BP neural network part of the process, all kinds of risk factors are default to the same degree of impact, without a rigorous distinction, which is adverse to the establishment of neural network model.

Considering the particularity of energy big data security and privacy risk, quantitative analysis method may not be able to reasonably determine the real impact degree of indexes. Therefore, AHP method is used to give weight to indexes in this paper, and various factors in complex problems are divided into interconnected and ordered levels to make them methodical. According to the subjective judgment structure of certain objective reality, the expert opinions and the objective judgment results of analysts are directly and effectively combined, and the importance of pairwise comparison of one level elements is quantitatively described.

Therefore, after the establishment of energy big data security privacy and risk assessment index system, according to the influence degree of each risk factors, the Delphi method is used to invite experts to quantify the importance between them, and the AHP method is used to give corresponding weights to 22 indexes.

- (1) Construct the judgment matrix.

The judgment matrix  $A = (a_{ij})_{n \times n}$  is established by pairwise comparison. In order to make the judgment quantitative, the quantitative scale is given for the evaluation of different situations. The scale specification is shown in Table 2.

- (2) Calculate the eigenvalue and eigenvector by the square root method and calculate the product of elements in each row of judgment matrix.

$$M_i = \prod_{j=1}^n a_{ij} \quad (i = 1, 2, \dots, n). \quad (2)$$

Calculate the  $n$ th root of  $M_i$ .

$$\bar{W}_i = \sqrt[n]{M_i} \quad (i = 1, 2, \dots, n). \quad (3)$$

Normalize the eigenvectors as the weight.

$$W_A = [W_1, W_2, \dots, W_n]^T. \quad (4)$$

Calculate the largest eigenvalue, where  $(AW)_i$  is the  $i$ th component of the vector  $AW$ .

$$\lambda_{\max} = \sum_{i=1}^n \frac{(AW)_i}{nW_i}. \quad (5)$$

- (3) Check for consistency.

The consistency index  $C.I.$  is

$$C.I. = \frac{\lambda_{\max} - n}{n - 1}. \quad (6)$$

Generally,  $C.I. \leq 0.10$  represents that the judgment matrix is consistent.

Obviously, with the increase of value  $n$ , the judgment error will increase, so the influence of  $n$  should be considered when judging the consistency, and the random consistency ratio  $C.R. = C.I./R.I.$  should be used, where  $R.I.$  is the average random consistency index. Table 3 shows the average random consistency index test values calculated by the judgment matrix.

**3.2. BP Neural Network.** BP neural network is a kind of multilayer neural network, which was proposed by Rumelhart in 1986. It is one of the most widely used neural network models at present. It can learn and store a large number of input-output pattern mapping relations. Its learning rule is to use the steepest descent method to continuously adjust the weights and thresholds of the network through back propagation, so as to minimize the mean squared errors of the network. It is usually composed of input layer, hidden layer, and output layer [23], and its network model is shown in Figure 2.

The basic unit of neural network is neuron. The principle formula is shown in formula (7); the commonly used activation functions are threshold function, sigmoid function, and hyperbolic tangent function. In formula (7), the input of neurons is represented by  $x_i$  ( $i = 1, 2, \dots, n$ ), the connection weights between neurons are represented by  $w_i$  ( $i = 1, 2, \dots, n$ ), the threshold of neurons is  $b$ , the activation function is  $f$ , and the output of neurons is  $y$ .

$$y = f\left(\sum_{i=1}^n x_i w_i + b\right). \quad (7)$$

For BP neural network, the mean square error  $E$  is often used as the index to judge the training performance of the model, shown in formula (8). The principle of minimizing the mean square error by adjusting the network weights is shown in formula (9), where  $e$  is the network error vector,  $y_i$  is the model output, and  $t_i$  is the target output.



TABLE 2: Scale specification.

Scale	Meaning ( $a_i$ vs $a_j$ )
1	The former is as important as the latter
3	The former is slightly more important than the latter
5	The former is obviously more important than the latter
7	The former is strongly more important than the latter
9	The former extremely is more important than the latter
2, 4, 6, and 8	The intermediate value of the above two adjacent judgments
The reciprocal of the above values	If the ratio of factors $i$ and $j$ is $a_{ij}$ , then the factor of the ratio of factors $j$ and $i$ is $a_{ji} = 1/a_{ij}$

TABLE 3: R.I. value.

Order number	1	2	3	4	5	6	7	8	9	10
R.I.	0	0	0.58	0.90	1.12	1.24	1.32	1.41	1.45	1.49

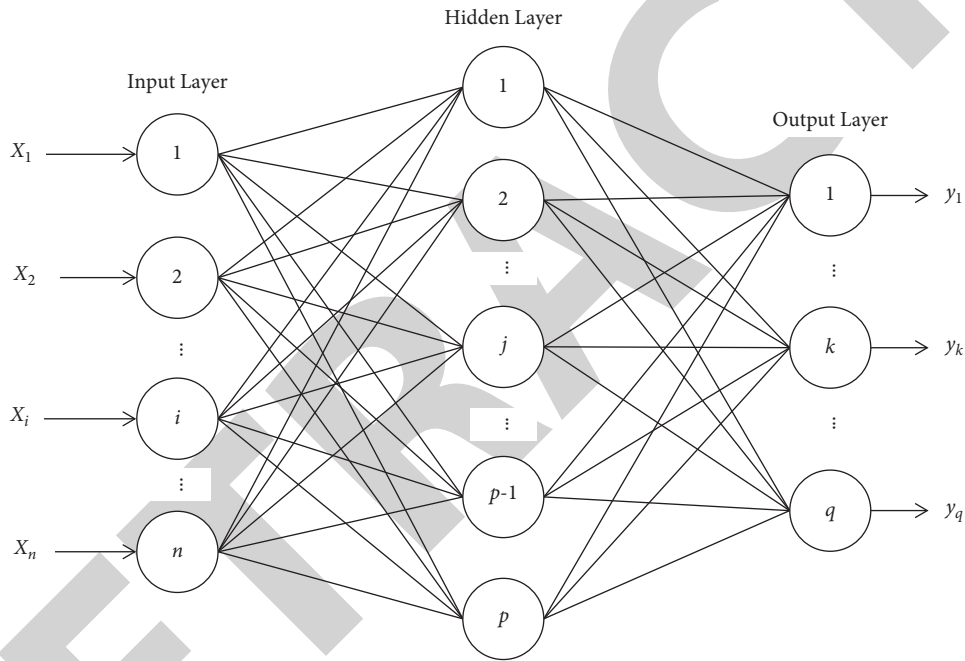


FIGURE 2: Structural model of neural network.

$$E = \frac{1}{n} \sum_{i=1}^n w_i (y_i - \hat{y}_i)^2. \quad (8)$$

$$\min E(e^T e) = \min E[(t - y)^T (t - y)]. \quad (9)$$

For the training model, the LM algorithm of neural network is used in this study. The basic method to reduce the error is as follows:

$$x(k+1) = x(k) - (J^T J + \mu J)^{-1} J^T e, \quad (10)$$

$$J^T J = H,$$

where  $H$  is the Jacobi matrix of the first derivative of the MSE function with respect to weights and thresholds.

**3.3. Genetic Algorithm.** Genetic algorithm (GA) is a computational model simulating the natural selection and genetic mechanism of Darwinian biological evolution theory. It is a method to search the optimal solution by simulating the natural evolution process [24].

Using genetic algorithm to get the optimal network weights and thresholds as the initial network weights and thresholds of the subsequent neural network model can not only overcome the defect that the traditional BP neural network is easy to fall into the local minimum, but also greatly improve the accuracy of model evaluation, so that the optimized BP neural network can better evaluate the samples. The elements of genetic algorithm include population initialization, fitness function, selection operator, crossover operator, and mutation operator.



Compared with binary coding, real coding can significantly reduce the length of coding and avoid the later decoding, with high accuracy. A series of parameters to be optimized, such as the connection weight, hidden layer node threshold, and output layer node threshold, are encoded by the  $s$ -order real matrix with the value range of  $[-1, 1]$ .

After coding, the selection, crossover, and mutation are performed. These three operations are based on the fitness value calculated by the fitness function as the assessment standard. The smaller the value, the larger the fitness value, and the better the individual. The fitness function of this study is the reciprocal of mean square error function, as follows:

$$F[f(x)] = \frac{1}{f(x)}. \quad (11)$$

In the selection operation, the most common roulette method is used. The probability of each individual being selected is positively proportional to its fitness value.  $N$  represents the population size,  $F_i$  represents the fitness function value of individual  $i$ , and  $p_i$  represents the probability of the  $i$ th individual being selected. The calculation way is as follows:

$$p_i = \frac{F_i}{\sum_{j=1}^N F_j}. \quad (12)$$

By using arithmetic crossover as formula (13), a new individual is obtained by using the linear combination between two individuals, where  $d$  is a random number uniformly distributed in  $[0, 1]$ :

$$\begin{aligned} c_1 &= p_1 * d + p_2 * (1 - d), \\ c_2 &= p_1 * (1 - d) + p_2 * d. \end{aligned} \quad (13)$$

Mutation operation refers to the random mutation of individual gene of the population, enhancing the local search ability of the algorithm and maintaining the diversity of individual population. The operation method of mutation of the  $j$  gene of the  $i$  individual  $a_{ij}$  is as follows:

$$a_{ij} = \begin{cases} a_{ij} + (a_{ij} - a_{\max}) * f(g), & r \geq 0.5, \\ a_{ij} + (a_{\min} - a_{ij}) * f(g), & r < 0.5, \end{cases} \quad (14)$$

where  $a_{\max}$  is the upper bound of gene  $a_{ij}$ ,  $a_{\min}$  is the lower bound of gene  $a_{ij}$ ,  $f(g) = r_2(1 - g/G_{\max})^2$ ,  $r_2$  is a random number,  $g$  is the current iteration number,  $G_{\max}$  is the maximum evolution number, and  $r$  is the random number of  $[0, 1]$  interval.

**3.4. Construction of AHP-GABP Model.** Compared with the traditional BP neural network, GABP model has a process of using genetic algorithm to optimize the weights and thresholds of the network, and this process can optimize the prediction performance of BP neural network to a certain extent. At the same time, using the AHP method to confirm the indicator weights can better define the importance of

indicators. The flowchart is shown in Figure 3. The steps to build the AHP-GABP model are as follows:

- (1) Use AHP method to process data.
- (2) Determine the topological structure of BP neural network.
- (3) After the weights are given by AHP, determine the input and output sample set and test sample set of training.
- (4) The network parameters to be optimized are real-coded to form their own chromosomes.
- (5) Determine the parameters of selection, crossover, and mutation.
- (6) Set the population size popu.
- (7) After inputting samples, each chromosome produces corresponding output after network transmission.
- (8) The fitness value of each chromosome is calculated by fitness function, and the selection operation is carried out according to the fitness value.
- (9) A new generation of population is generated by crossover and mutation.
- (10) Repeat steps 6–8 until the fitness value of the optimal individual and the fitness value of the population do not rise within the specified number interval, or the fitness value of the optimal individual reaches the set threshold, or the number of iterations reaches the algebra set in advance, the algorithm stops, and the optimized network parameters are obtained.

## 4. Evaluation Process

### 4.1. Model Training

#### 4.1.1. Network Design

(1) *Network Structure Determination.* The paper selects 22 assessment indexes to assess the security and privacy risk of energy big data, so the number of input layer nodes is 22. In general, if the number of hidden layers is more, the error of assessment results will be smaller, but it will also bring the disadvantages of network complexity, thus reducing the efficiency of training [25]. For the multi-input single-output network model established in this paper, in order to increase the approximation effect and convergence, and reduce the oscillation in the simulation process, the number of hidden layer nodes is determined by referring to equation (15) and combining with the actual simulation results.

$$S_1 = \sqrt{m + n} + a, \quad (15)$$

where  $m$  represents the number of input layer nodes,  $n$  represents the number of output layer nodes,  $a$  takes a random integer between 1 and 10, and  $S_1 = 12$  is determined after trial calculation. The final MATLAB structure is shown in Figure 4.

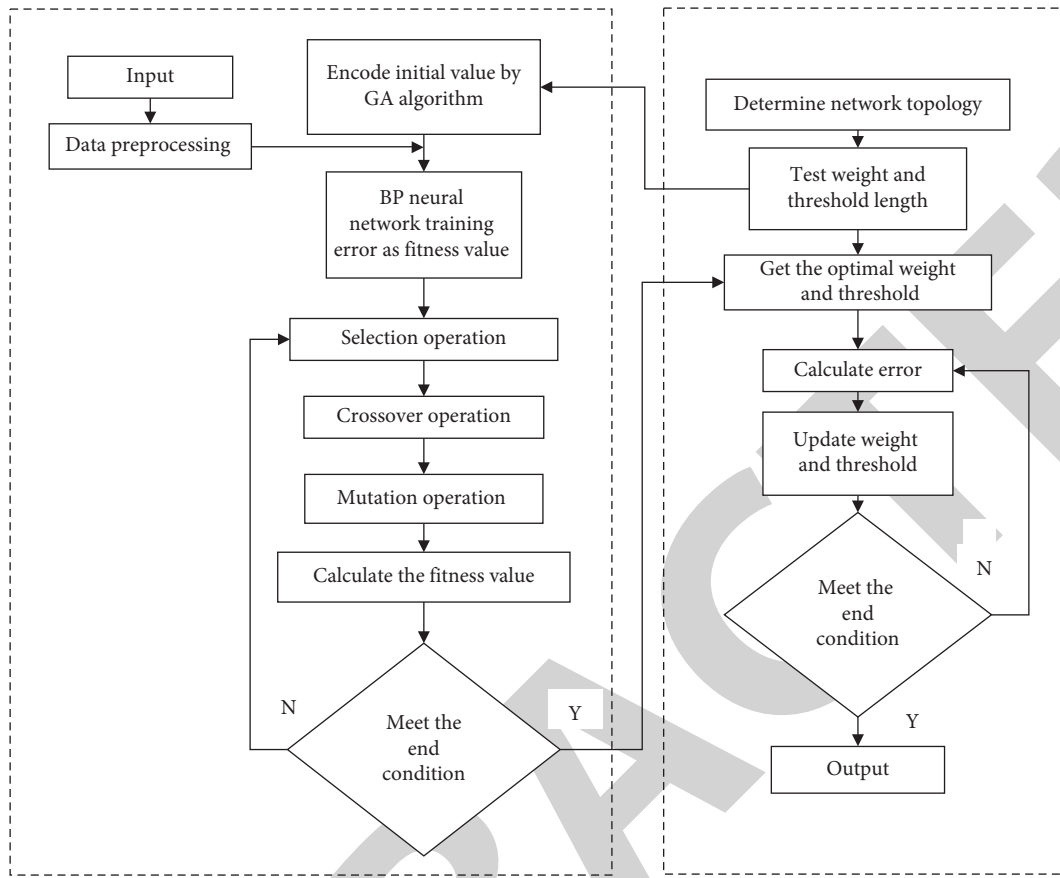


FIGURE 3: AHP-GABP flowchart.

(2) *Parameter Setting.* This study uses feedforward net to create function, trainlm to train function, logsig to transfer function, sigmoid to activate function, and MSE to express error  $E$ . The training times is 100, the learning rate is 0.01, and the training error target is 0.01. For the part of genetic algorithm, the number of population is set to 100, the maximum evolution algebra is set to 100, the variable precision is  $1e - 6$ , the crossover probability is 0.8, and the mutation probability is 0.2.

4.1.2. *Training Results.* After reading the literature and cases about the security and privacy risk of energy big data, a total of 44 samples are collected, including 36 training samples and 8 test samples. Some of the training data are shown in Table 4. The model training is realized by MATLAB programming and the development of Goat genetic algorithm toolbox.

The training data is input into the program, and the convergence curve of genetic algorithm optimized BP neural network is shown in Figure 5. It can be seen from the figure that the BP neural network algorithm after genetic algorithm optimization finds an optimal path optimal solution when the population iteration is about 60 generations, which shows the superiority of genetic algorithm in optimizing the weight and threshold of BP neural network. It can also be seen that the optimal function tends to be stable when the iteration reaches nearly 70 generations.

The BP neural network and the optimized genetic BP neural network are compared, and their error values are calculated. The final experimental results are shown in Table 5. Through analysis and comparison, in 8 groups of test samples, AHP-GABP prediction has significant advantages over BP prediction, with smaller error, shorter evaluation cycle, and greater improvement in evaluation performance. As shown in Table 5 and Figure 6, the BP neural network optimized by genetic algorithm improves the shortcomings of BP neural network, thus greatly improving the predictability of neural network. At the same time, the application assessment results of the BP neural network optimized by genetic algorithm in the energy big data security and privacy risk are basically consistent with the actual expert assessment results, which proves that the training network has high accuracy.

## 4.2. Model Applications

4.2.1. *Background.* Z power grid system uses its energy big data information to provide data services related to economic development. It can provide more reliable data support for poverty alleviation effect evaluation, credit evaluation, census, pollution monitoring, and work resumption evaluation. According to the energy big data security and privacy risk assessment index system designed above, the complete evaluation steps of big data security and privacy risk of this power grid system are as follows:

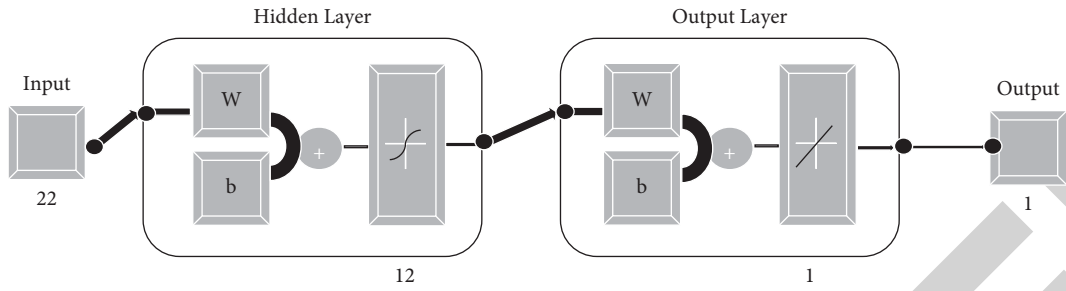


FIGURE 4: MATLAB structure.

TABLE 4: Training samples.

	1	2	3	4	5	6	7	8
A1	0.7028	0.5020	0.6024	0.6526	0.5522	0.6024	0.6024	0.5522
A2	0.3784	0.8514	0.9933	1.0406	0.6622	0.6622	0.6149	0.2838
A3	0.5400	0.2700	0.7200	0.8100	0.4050	0.2700	0.4500	0.1800
A4	0.3612	0.3010	0.4515	0.5418	0.2408	0.3010	0.3311	0.2408
B1	0.0516	0.0344	0.1290	0.1290	0.0688	0.0774	0.0688	0.0430
B2	0.0748	0.0408	0.1088	0.1360	0.0680	0.0816	0.0476	0.0476
B3	1.4922	0.9948	1.9067	0.9948	1.4922	1.2435	1.0777	0.9948
B4	1.9100	2.1965	2.0055	1.1460	2.0055	1.2415	1.3370	0.9550
B5	0.4992	0.6240	0.5408	0.3328	0.5824	0.2496	0.4160	0.2496
C1	0.1854	0.4326	0.4326	0.1545	0.5871	0.3090	0.2781	0.2163
C2	0.0648	0.1620	0.1944	0.0810	0.2430	0.1134	0.1134	0.0486
C3	1.6300	1.6300	2.2820	1.6300	1.9560	1.7930	1.6300	1.4670
C4	1.8032	1.0304	2.0608	0.7728	1.2880	2.7048	1.8032	1.2880
C5	0.5130	0.5130	1.0260	0.4617	0.5130	0.9747	0.4104	0.4104
D1	0.2808	0.5967	0.6318	0.2808	0.3159	0.5265	0.2106	0.3159
D2	0.5136	0.3531	0.6741	0.3210	0.3210	0.6420	0.2247	0.2247
D3	0.1932	0.1380	0.2208	0.0966	0.0828	0.2346	0.0966	0.0828
D4	0.3108	0.4144	0.3626	0.2331	0.1813	0.4662	0.2072	0.2072
D5	0.2100	0.2520	0.1680	0.1260	0.1680	0.3360	0.1260	0.1680
E1	0.7434	0.2891	0.7847	0.4956	0.4130	0.5369	0.8260	0.4130
E2	0.1261	0.1164	0.1358	0.0970	0.0582	0.1358	0.1552	0.0582
E3	0.0156	0.0117	0.0312	0.0156	0.0195	0.0234	0.0546	0.0234
Output	0.7813	0.7558	0.9149	0.6012	0.7801	0.7889	0.5989	0.4569
Risk level	4	4	5	4	4	4	3	3

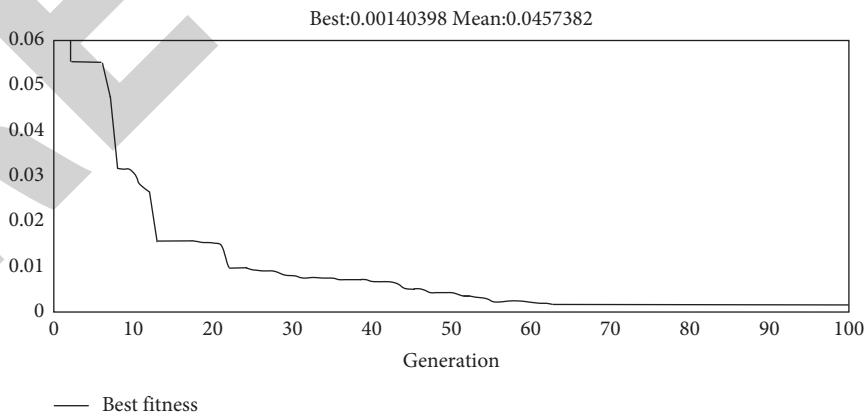


FIGURE 5: Fitness value.

TABLE 5: Comparison of training sample error between BP neural network and AHP-GABP neural network.

Sample number	Real value	Predictive value		Error	
		BP	AHP-GABP	BP	AHP-GABP
1	0.7813	1.4328	0.7717	0.6515	-0.0096
2	0.7558	0.4292	0.7443	-0.3266	-0.0115
3	0.9149	1.9229	0.8984	1.0080	-0.0165
4	0.6012	-1.0077	0.5324	-1.6089	-0.0688
5	0.7801	0.3213	0.7503	-0.4588	-0.0298
6	0.7889	0.7989	0.8125	0.0100	0.0236
7	0.5989	1.3607	0.5810	0.7618	-0.0179
8	0.4569	-0.0912	0.4442	-0.5481	-0.0127

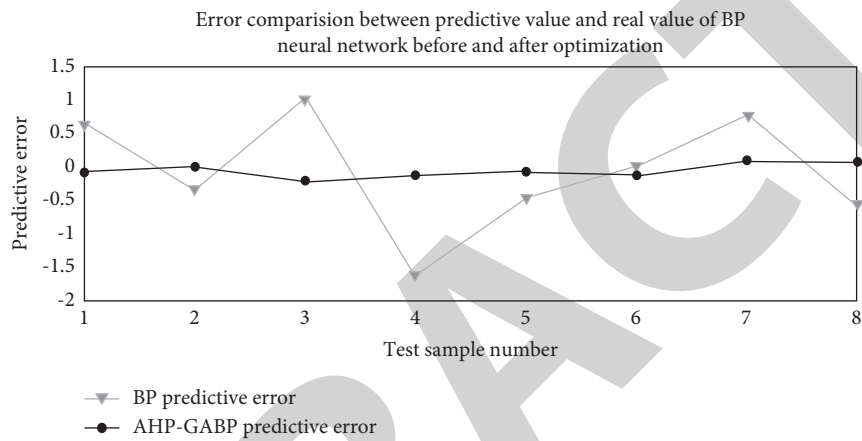


FIGURE 6: Error comparison between predictive value and real value.

TABLE 6: Weights of risk assessment indexes.

First-level index	Second-level index	Weight
Data collection A	Software and hardware fault risk A1	0.0601
	Damage or consumption risk of energy infrastructure A2	0.0270
	External malicious attack A3	0.1058
	Irresistible force risk A4	0.0108
Data transmission B	Malicious intercepting risk B1	0.1774
	Malicious tampering risk B2	0.1133
	Data distortion risk B3	0.0720
	Access control risk B4	0.0210
	Cloud platform risk B5	0.0317
Data storage C	Sleeping data risk C1	0.0124
	Quality of data input risk C2	0.0163
	Data leakage risk C3	0.1237
	Management data destruction risk C4	0.0324
	Virus intrusion risk C5	0.0680
Data use D	Multi source data fusion risk D1	0.0432
	Privacy awareness of business personnel D2	0.0046
	Data parsing risk D3	0.0113
	Data regulatory risk D4	0.0165
	Manage authorization risk D5	0.0071
Data destruction E	Data residual risk E1	0.0135
	Data backup risk E2	0.0245
	Termination of cloud service agreement risk E3	0.0074

TABLE 7: Case assessment result.

Sample	AHP-GABP result	Risk level
1	0.1710	1
2	0.1715	1
3	0.1033	1

- (i) Calculate the index weights using AHP method.
- (ii) Collect relevant data of this grid system, invite relevant department heads to score the 22 risk assessment indicators, and standardize the data with the weights as the input values of the AHP-GABP model.
- (iii) Use the above trained AHP-GABP network model; the output values are evaluated, and the risk level is defined according to the risk classification method.

**4.2.2. Initial Index Weight of AHP Method.** In this study, AHP method is used to assign weights to the primary and secondary indexes, respectively. After the consistency check, the final weights of 22 indexes are obtained as shown in Table 6.

**4.2.3. Assessment Results.** In this study, three groups of relevant data collected by the power grid system are selected. After training, the AHP-GABP neural network model is established. Firstly, it is necessary to verify whether the evaluation model is reasonable. Secondly, it is necessary to assess the risk. The assessment results are shown in Table 7, which shows that the risk level of the power grid system is class 1, which is similar to the conventional risk performance of the power grid system. The risk level is low, and there is no need to do special treatment, and regular inspection should be done. It also shows that the AHP-GABP algorithm is reasonable and correct in the evaluation and prediction, with high prediction accuracy, objective and fair evaluation results, wide application range, and high practical application value.

## 5. Conclusion and Development Suggestions

To sum up, in the process of controlling the energy big data security and privacy risk, the risk of each stage cannot be ignored. On the premise of comprehensively considering the cloud environment and risk factors, this paper divides the potential energy big data security and privacy risk of each stage as comprehensively as possible according to the life cycle of big data, and uses AHP method to allocate weights for the indexes, which provides a reference for the future energy big data research. At the same time, this paper optimizes the BP neural network model based on the evaluation, and tries to apply the AHP-GABP method to the risk evaluation of energy big data security and privacy, which greatly reduces the risk that the random selection of initial weights and thresholds in BP algorithm leads to the model training easily falling into the local minimum, and improves the accuracy of neural network model assessment and predication and realizes the application of AI related knowledge in the field of energy.

The AHP-GABP model is applied to evaluate the security and privacy of the energy big data, and the evaluation results are good. According to the case and expert interviews, the following development suggestions are summarized for the common risks of energy big data security and privacy.

**5.1. Pay Attention to the Security of the Whole Life Cycle of Energy Big Data.** Energy big data comes from production data and operation and management data, and its protection should focus on the whole life cycle of data collection, transmission, storage, use, and destruction. From policy and system requirements to technical management and control, we should comprehensively assess the threat exposure of critical data and make targeted protection strategies at all stages to ensure the security of core data assets.

**5.2. Strengthen Technical Protection of Energy Industry Based on Big Data Security.** The energy industry should establish a comprehensive threat early warning technology based on security big data, break through the traditional mode, and more actively detect potential security threats. The introduction of big data analysis technology in threat detection can more comprehensively detect attacks on data assets, software assets, physical assets, personnel assets, service assets, and other intangible assets supporting business [26]. At the same time, the scope of the analysis content can be expanded. The threat analysis window can span several years of data, so the threat detection ability is stronger and can effectively respond to the attack [27].

**5.3. Consider Security and Privacy Issues from a Strategic and Long-Term Perspective.** Big data brings opportunities and challenges to the energy industry. The more widely it is applied, the greater the value it brings. The concept of security management centered on data security will change the traditional working ideas [28]. We must recognize the new changes, new features, and new trends of big data security, and deeply analyze the outstanding problems existing in big data security under the current situation. In order to ensure that the development strategy of energy big data information security is consistent with the national conditions and constantly improves, it is necessary to plan the key layout of big data application, key technology research and development, data protection, laws and regulations.

With the rapid development of cloud computing and the continuous improvement of digital level, the energy big data security and privacy risk evaluation index system can be further improved. At the same time, with the enrichment of data indicators and training models, the model proposed in this paper can also be better optimized and expanded to other fields for more accurate evaluation and prediction in the future.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.